LOCAL ADAPTATION AND THE GENETIC BASIS OF ADAPTIVE VARIATION IN WILD PLANTS

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by
BILLIE ALETHEA GOULD
May 2013

LOCAL ADAPTATION AND THE GENETIC BASIS OF ADAPTIVE VARIATION IN WILD PLANTS

Billie Alethea Gould, Ph. D.
Cornell University 2013

The phenomenon of local adaptation is a key player in the evolution of plants, which are shaped by their environments perhaps more than any other organismal group. Botanists have often demonstrated adaptive trait differences between plant populations in different environments, and the concept of the "ecotype" was itself coined by an early plant biologist (Turesson, 1922). We continue striving to understand how the push and pull of selection and gene flow across heterogeneous environments contributes to the maintenance of genetic diversity and influences fundamental aspects of species biology such as geographic distribution, morphological diversity, and population response to environmental change. Understanding how often and why local adaptation occurs is an area of research that links both basic and applied branches of plant biology.

In my dissertation I address fundamental questions regarding local adaptation in plants including: 1) What is the prevalence and role of local adaptation in determining the geographic distribution of a species; 2) What are the genes involved in local adaptation in the wild; and 3) do adaptive phenotypes evolve by similar genetic pathways in related species with different histories of selection. To address the first question, I examined the role of trait differentiation in range boundary formation in the annual wildflower *Clarkia xantiana ssp. xantiana*. To explore the second and third questions, I examined patterns of local adaptation and its genetic basis in the temperate grass *Anthoxanthum odoratum.* I

focused on the ecologically and agriculturally important trait of tolerance to aluminum in acid soils.

In *C. x.* ssp. *xantiana* I have shown that adaptive differentiation between populations is common. Counter to theoretical expectations however, local adaptation to conditions at the range edge does not preclude the existence of substantial heritable trait variation there and is thus unlikely to restrict adaptation to conditions beyond the range edge. In *A. odoratum* local adaptation is also prevalent even at the small spatial scale of experimental plots within a single hay meadow at the long-term ecological Park Grass Experiment. Using genomic techniques, I demonstrate that adaptation to soil Al stress in this wild grass has many genetic similarities to cultivated grasses, but also likely involves previously undescribed genetic pathways. Both novel and canonical pathways are also likely to have been the targets of selection during the process of local adaptation during the history of the experiment. In combination, these studies reaffirm the prevalence of local adaptation in nature, but they also demonstrate that simple theoretical predictions about the exitence of local adaptation, its genetic basis, and its ecological consequences are suspect. Direct studies of adaptive traits and their underlying genes in diverse organisms will continue to be critical for understanding the true nature of the complex interaction between selection, gene flow, and genetic architechture to produce what we observe in the natural world.

BIOGRAPHICAL SKETCH

The author was born to Carol and Michael Gould in March 1980 and grew up in the small town of Harwinton, CT. She had an early love of wild plants which was nurtured though the encouragement of her parents and many excellent teachers along the way. She attended Brandeis University earning a B.S. in Biology with a minor in Environmental Studies. She interned with the Environmental Protection Agency and volunteered with the New England Wildflower Society until in 2004 when she became a research assistant with Dr. Elena Kramer at Havard University. During this formative time she worked on the developmental genetics of the columbine plant (*Aquilegia*) and gained a love of genetics as it relates to plant diversity. In 2007 she began graduate studies at Cornell University with the goal of combining her interests in plant genetics and the environment. She has completed the present work with the guidance of her major advisor, Dr. Monica Geber in the Department of Ecology and Evolutionary Biology. At the time of this writing, the author has started post-doc work continuing to investigate the genetics of plant local adaptation at the University of Toronto.

Dedicated to my parents Carol and Michael Gould

Thank you for believing in me,

and for always being there

rain or shine.

ACKNOWLEDGMENTS

First and foremost I would like to thank Monica Geber my advisor without whom this work would not have been possible. Her willingness to accept me as a student and subsequent compassionate guidance has encouraged me to to think broadly, to believe I could achieve challenging research goals, and has helped make me the scientist I am today. I would also like to thank Susan McCouch who so kindly took me into her lab and has provided invaluable support and enthusiasm for this work since its inception. Thank you also to the other members of my committee Rick Harrison and Kelly Zamudio for their thoughtful comments throughout the process.

I would also not have been able to complete this work without the moral and scientific support of so many wonderful friends and collegues in the Department of EEB who have provided insightful comments and support. This includes especially my fellow EEB grads Alexis Erwin, Renee Petipas, Nancy Chen, Erica Larson, Indrani Singh, Susan Cook-Patton, Marjorie Weber, Jesse Bellemare, Jill Anderson, Sarah Reily, Laurie Evanhoe, and Brian Barringer. I would also like to thank my collegues outside of EEB that so generously provided their time and advice on technical aspects including Paul Cooper, Steve Bogdanowicz, Chih-Wei Tung, Adam Famoso, Moni Singh, Jianyong Yu, Eric Craft, John Shaff, Denise Costich, Qi Sun CBSU staff, and Julie Oh.

Lastly, I would like to express appreciation for the financial research support I have received as a graduate studnet. A large protion of funding came from the Cornell Center for Comparative and Population Genomics. Additional funding came from an NSF IGERT grant, the Keicheffer Adirondack Research fund, and the Cornell Chapter of Sigma Xi. In addition an NSF Graduate Research Fellowship enormously facilitated my completion of this work. Financial support for new graduate student work provided by these institutions has been vital for both my development as a scientist and the realization of this new body of ecological genetic research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Adaptive differentiation between central and range edge populations of *Clarkia xantiana* ssp. *xantiana*[1]

Abstract

Theory predicts that lack of heritable trait variation and/or maladaptive gene flow has the potential to promote the formation of species geographic range boundaries even in the absence of barriers to dispersal. However, little is known about the patterns and drivers of differentiation across species' ranges and whether they influence boundary formation in the field. Using field measurements, two common garden studies, and Q_{ST}-F_{ST} analyses, we examined the environmental and genetic influences on plant phenotype across the geographic range of *Clarkia xantiana* ssp. *xantiana*. *Clarkia* is an annual plant endemic to California whose eastern range border occurs without apparent physical barriers to dispersal. We detected two strong but opposing environmental gradients across the range that likely affect the evolution of plant traits. Both phenotypic and genetic trait differentiation among populations were significant and correlated with environmental gradients. Q_{ST} - F_{ST} comparisons suggest local adaptation for flowering time, the degree of branching, and herkogamy. Populations at the range-edge were phenotypically and genetically differentiated from central ones. However, the amount of quantitative genetic variation within populations did not differ with proximity to the range edge. Traits were generally heritable within populations and not tightly correlated. Our results collectively suggest that despite significant local adaptation across the range, there is little

[1] Co-authors on this work: Dave Moeller, Vince Eckhart, Peter Tiffin, and Monica Geber.

evidence that maladaptive gene flow or lack of heritable trait variation at the range edge strongly inhibits adaptation and population expansion beyond the eastern range border.

Introduction

Understanding the factors that determine species' geographic range limits is a long-standing problem in ecology and evolution (Grinnell, 1917; MacArthur, 1972; Brown & Lomolino, 1998; Gaston, 2003; Geber, 2011). Why, with ample time for the forces of mutation and selection to produce adaptation, do most native species occupy stable niches at their range boundaries? As Mayr (1963) aptly put it, "One would expect that a few individuals would survive in a zone immediately outside the species border and form a new local population which becomes gradually better adapted . . . . the species range to grow by a process of annual accretion like the rings of a tree" (p.524). Since Mayr's time, studies have demonstrated that, from a demographic standpoint, range limits form where death rates exceed birth rates and where extinction rates exceed colonization rates. From an evolutionary standpoint, populations fail to persist in marginal environments because adaptation fails to occur in response to some (suite of) novel environmental factor(s) (Kirkpatrick & Barton, 1997; Bridle & Vines, 2007; Geber, 2011).

Theoretical work has modeled how demographic and evolutionary factors might cause the formation of stable range limits, but few species have been investigated in enough detail to distinguish among competing ecological and evolutionary mechanisms. On one hand, peripheral populations may be demographically unstable and maintained largely by immigration. For example, source-sink or colonization-extinction dynamics from the range center to periphery can allow edge populations to persist, despite high extinction rates in the absence of immigration (Pulliam, 1988; Hoffmann & Blows, 1994; Kirkpatrick & Barton, 1997). In this case, peripheral populations may show limited adaptation to range-edge environments, and limited adaptive

differentiation from central populations because of continuous introduction of "centrally-adapted" alleles. On the other hand, peripheral populations may be stable without immigration, and are expected to exhibit adaptive differentiation from central ones. Examining patterns of population differentiation in ecologically important traits from the center to the periphery of geographic ranges can provide particularly useful insight for differentiating between these possibilities.

For stable peripheral populations, theory suggests that a lack of genetic variation underlying ecologically important traits – due to strong selection, small population size, and/or inbreeding – can hamper adaptive evolution to conditions beyond the range limit. Alternatively, persistent gene flow from central to unstable peripheral populations, while limiting adaptive differentiation, is likely to maintain genetic variation in edge populations (Hoffmann & Blows, 1994). Some studies have investigated population genetic structure and variation among central and marginal populations of plant species (Eckert, Samis, and Lougheed, 2008; Platt *et al.*, 2010; Moeller *et al.*, 2011; Keller *et al.*, 2011), but little is understood about the level of heritable trait variation and its potential correlation with environment. Few studies of plants have examined the relationship between environment, phenotypic differentiation, and heritability in combination from the center to margin of a species' range.

Here, we use a field study of natural populations combined with both a population-structured and a family-structured common garden study to examine patterns of quantitative trait variation both within and among populations that span the majority (the "core") of the geographic range of *Clarkia xantiana* ssp. *xantiana* (Onagraceae). *Clarkia x.* ssp. *xantiana* is an outcrossing annual plant that occupies a narrow endemic range in the mountains of inland central and southern California (Lewis & Lewis, 1955; Eckhart & Geber, 1999). In the Sierra Nevada,

the core of its geographic range, it spans a range of elevations where soils, temperature, precipitation, and topography vary (Eckhart *et al.*, 2010, 2011). The range edge in the eastern part of the mountains has no known physical barrier to dispersal. Demographic work and reciprocal transplant experiments have demonstrated that habitat favorability is low beyond the eastern border (Geber & Eckhart, 2005; Eckhart *et al.*, 2011). Central populations are somewhat more genetically diverse at putatively neutral loci and there is generally more gene flow from central to edge populations than vice versa (Moeller *et al.*, 2011).

We used field and common garden experiments to measure phenotypic differentiation among populations spanning from the range center to the range edge, and tested whether differences between populations are likely to be caused by local adaptation (rather than purely by drift). This is a prerequisite for maladaptive gene flow to cause the formation of a range boundary. We compared patterns of phenotypic differentiation in the field and common garden to variation in environmental variables known to affect plant growth, survival and reproduction. If trait differentiation in the field – especially differentiation that is correlated with environment – disappears when populations are raised in a common garden, then it is likely that phenotypic plasticity explains the variation we observe in the wild. If phenotypic differentiation observed in the field is still apparent in a common garden and is correlated with the field environment, trait differences are likely to reflect genetic trait divergence in response to those environmental factors. To further test whether trait divergence is adaptive, we used a family-structured common garden to compare heritable trait differentiation between populations (Q-statistics, (Spitze, 1993; Whitlock, 2008; Keller *et al.* 2011) to between population differentiation in neutral markers (F-statistics, [Wright, 1951; Hudson, Boos, and Kaplan, 1992]) spanning from the range center to the range edge. The family-structured common garden experiment also allowed us to directly

assess levels of quantitative genetic variation within populations and ask whether peripheral populations harbored less variation than central populations.

Materials and Methods

Study Species: *Clarkia xantiana* ssp. *xantiana* is an outcrossing winter annual that occurs in grassland, pine-oak woodland, and chaparral habitats. It occupies a narrow geographic range, across approximately 4,000 km^2 in the mountains of inland central and southern California (Fig. 1a; Lewis & Lewis, 1955; Eckhart & Geber, 1999). Populations are mostly discrete, separated by a few hundred meters up to tens of kilometers. Plants are pollinated by both generalist and specialist bees (Moeller, 2006; Moeller, Geber, and Tiffin, 2012), and seeds are gravity dispersed over short distances.

We studied *C. xantiana ssp xantiana* in the core of its geographic range in the southern Sierra Nevada (Fig. 1A). In this area the southwestern range edge occurs where the environment is heavily altered by human development and habitat changes abruptly. Eastward from there, populations occur along the steep rise into the mountains, on the slopes of the Kern River Canyon above the western Kern Valley; and northward along the North Fork of the Kern River, which parallels the eastern range edge (Eckhart & Geber, 1999). There is a gradient of declining temperature and increasing spring precipitation from southwest to northeast across this portion of the range (Eckhart et al. 2011). A genealogically distinct self-fertilizing subspecies, *C. x.* ssp. *parviflora,* is parapatric with *C. x.* ssp. *xantiana* along the latter's eastern range edge (Pettengill & Moeller, 2012a). The subspecies are in secondary contact and hybridization is limited (Pettengill & Moeller, 2012b). Hereafter in the text the term '*C. xantiana*' will refer to the *xantiana* subspecies only, unless otherwise noted.

Environmental measurements: In the spring of 2008, we studied 30 populations of *C. xantiana* distributed across its core range in the southern Sierra Nevada (Fig. 1a). The geographic location of each population, in units of easting (longitude) and northing (latitude) (UTM NAD 27, zone 11 north) was recorded with GPS.

We measured five environmental variables at each site: average daily temperature and total precipitation between February and June (the spring growing season), slope, aspect, and solar radiation at the spring equinox. Temperature and precipitation were measured using a network of 20 weather stations distributed across the range. For sites lacking weather stations, we used the Spatial Analyst Extension of ArcGIS 9.3 to estimate temperature and precipitation through spatial interpolation (see Eckhart *et al*., 2011). Slope was calculated as the average of three measures taken at separate locations within a site. Aspect was measured as the midpoint of compass values from two locations within a site and converted to linear azimuth, a measure of the north-south orientation (Warren, 2008). Solar radiation (predicted for the spring equinox) was estimated in ArcGIS 9.3 in units of Watt-hours per m^2, under a 12-hour day.

Field trait measurements: In each of the 30 *Clarkia* populations we measured phenotypes on 61-149 plants (average = 111), along 2 or 3 parallel transects spanning each site. Plants were measured every 5 paces, with more plants sampled from larger populations. Along transects we also recorded the number of plants with evidence of herbivore browsing damage, which primarily comes from mammal browsing and to a smaller extent from insects. (We took phenotypic measurements only on plants that lacked heavy browsing damage.) For each plant, we measured height (cm) and number of primary branches. For plants that had at least one

branch, we calculated an index of 'branchiness' as the total number of branches divided by plant height. We also calculated the total percentage of plants with and without branches among all the plants measured at each site. We counted the total number of flowers on each plant and recorded the maturation stage of each flower. *Clarkia* flowers are protandrous, progressing through 5 stages: (1) developing bud, (2) male, when anthers are fully dehiscent, (3) female, when the stigma is receptive, (4) green immature fruit, and (5) brown dehiscent fruit. The flowering stage for each plant was calculated as the average maturation stage of all flowers on the plant, and then the population average flowering time was calculated as the mean of these values across all plants measured at the site. By this measure, populations with a higher value flowered earlier in the season. Because measuring all 30 populations took 11 days in the field, we adjusted flowering time for differences in sampling date by re-measuring the flowering stage of 13 populations (76-139 plants/population) a second time, 5-14 days following the initial measurement. We calculated the change in average flowering stage per day and used this rate of change to standardize flowering stage to a single date at the middle of the measurement period. For populations that were not re-measured, the average population change (0.104 stages/day, stdev = 0.029) was used for standardization.

We collected one fruit from each of 15 plants per population, weighed four seeds per fruit to calculate an average seed weight. The seeds were then used in a 30-population common garden greenhouse experiment to compare phenotypic traits of the populations under uniform conditions (see below). The census population size of each population was estimated as the product of the average number of fruiting plants across 50-100 haphazardly distributed sampling plots (0.5 m^2) and the site area obtained from GPS coordinates of the circumference of each site.

30 population common garden: To assess whether phenotypic differences among field populations persist in a common environment, we planted four seeds from each field-collected plant into 16 cm 'cone-tainers' (Suewe & Sons, Tangent, OR) in MetroMix 360 potting soil (Sun Gro, Bellevue, WA) (Fig. 1b). Pots from each plant and source population were randomly assigned to seven blocks in a growth chamber under short 10-hour days at cool temperatures (12/10°C day/night) for four weeks to simulate winter conditions preceding germination. Daytime temperature was then increased to 18°C for an additional four weeks to simulate spring warming. We thinned out all but the first emerging seedling from each pot, randomized the pots into six complete spatial blocks in the greenhouse, with two plants per population per block, and grew them at 20/18°C, 12-hour days under supplemental metal halide lighting. For each plant we measured the date of first flower and on the same day we measured the height, the number of primary branches (from which we calculated the index of branchiness), and the number of flowers (including buds over 0.5 cm long). Because of incomplete germination in some populations and some adult plant mortality, we obtained vegetative measurements on an average of 11.4 plants per population and floral measurements on an average of 11.3 plants per population (total N = 341 plants).

Family-structured common garden: For six of the 30 populations, we generated full-sib families of seeds in the greenhouse. Two populations were used from each of three geographic regions: the range center, a region intermediate between the center and the eastern edge, and the eastern edge (Fig. 1a). For each population, a single seed was germinated in a growth chamber (as above) from each of 48 field collected plants, and seedlings were paired at random (within populations) and crossed reciprocally at flowering across six temporal blocks in the greenhouse

to generate 24 full-sib families per population. We measured the average weight of eight seeds per family (four from each reciprocal parent), planted seeds in cone-tainers, and allowed germination in the growth chamber (as above). We transferred the seedlings into each of four spatial blocks in the greenhouse. This was done for two separate cohorts of plants (temporal blocks), together totaling 576 plants. For each plant we measured the date of first flowering, and on the same day we measured plant height and the number of primary branches from which we calculated branchiness. We also measured two floral characters related to the degree of outcrossing in this subspecies: herkogamy (the distance between anthers and the mature stigma), and the length of the apical petal (a measure of flower size) on the second flower to open on each plant (Moore & Lewis, 1965; Runions & Geber, 2000; Moeller & Geber, 2005).

Statistical analyses

Relationships between plant traits and the environment: To examine the relationship between the five environmental variables and plant traits, we used a stepwise regression to determine the best predictors of each plant trait based on the corrected minimum Akaike information criterion (AICc; Hurvich & Tsai, 1989). Significant predictors for each trait (at $p \leq 0.05$) were retained in the final model, and where more than one environmental variable was included in the best model we used partial linear regression to examine the relationship between the plant trait and each environmental variable separately, holding the other environmental variables constant. We computed the residuals of the plant trait from a regression model that included all significant environmental predictors except the focal predictor, as well as the residuals of the focal environmental variable from a regression against the other significant environmental predictors. We then plotted the residual trait values against the residuals of the

environmental predictor (Fig. S3) for all populations, both in the field and in the common garden.

Variance components, heritability and trait correlations: For the field study, 30-population common garden study, and the family-structured study, we analyzed sources of phenotypic variation for each trait using general linear random effects models using the *nlme* and *lme4* packages in R (R Development Core Team, 2011). Average plant height, branchiness, flower number, and seed weight were log-transformed to improve normality.

For the field data, population was included in the general linear model as the only random factor, and the significance of the population effect was evaluated with a chi-square test (with 1 df) of twice the difference in the log-likelihood of models with vs. without the population. Estimates of phenotypic variances within (residual) and among populations for each trait were obtained using restricted maximum likelihood with the *lmer* function in R.

For the greenhouse data, we first used a model that included greenhouse block as a random factor, but dropped the block effect for all traits except the weight of greenhouse-generated seed (family-structured study) because it was not statistically significant. For seed weight, we used the residual values (from a model that included only greenhouse block) in subsequent analyses of population and family sources of phenotypic variation.

Population was included as a factor in analyses of both common garden studies, and family nested within population was included as a factor in the family-structured study. Both were included as random effects. The significance of population and family effects was tested by comparing the log-likelihood of models with vs. without the effect in question (using *lme* and maximum likelihood estimation). Final estimates of variance components (among vs. within

populations) for the common garden study of 30 populations and for the family-structured study (among population, among family and within family variances) were estimated using restricted maximum likelihood with the *lmer* function. Population mean phenotypes for each trait in each data set were estimated as the random effects coefficients of populations in the general linear models.

We estimated the upper bound on broad sense heritability in the family-structured study as twice the intra-class correlation ($H^2 = 2*(V_{\text{between family}} / (V_{\text{between family}} + V_{\text{within family}}))$) for each trait (V=phenotypic variance; Falconer, 1960). Standard errors were calculated according to (Roff, 1997). Within- and between-family components of trait variation were obtained from general linear random effects models, with family as a random effect, for each population separately. For each of the six populations, we also estimated pairwise trait genetic correlations as the standard product moment correlation between family means for each trait (Via, 1984).

***F* and *Q* Statistics:** F_{ST} was calculated from nucleotide variation at eight nuclear loci, each sequenced for 135 individuals (20-23 individuals per population, further details in Moeller *et al*., 2011). These loci were chosen on the basis of single copy status in the genome and successful PCR amplification rather than level of sequence polymorphism and are putatively neutral with respect to recent or ongoing selection. They together provide an estimate of background genome-wide differentiation and are hereafter referred to as 'neutral' loci. We examined differentiation between each pair of populations (Hudson *et al.*, 1992) and hierarchically among geographic regions (F_{CT}; regions: center, intermediate, and edge), and between populations nested within regions (F_{SC}), using AMOVA (Excoffier, Smouse, and Quattro, 1992).

Q_{ST} is a measure of population differentiation in quantitative genetic traits that is directly

comparable with the neutral genetic measure of differentiation, F_{ST} (Spitze, 1993, Leinonen & Merilla, 2008). For each trait, we calculated an overall (broadsense) Q_{ST}, two hierarchical Q_{ST}'s (among regions (Q_{CT}); and between populations within regions (Q_{SC})), as well as a Q_{ST} for each pair of populations. These statistics were calculated as:

Q_{ST} (overall, among populations): $V_{pop} / (V_{pop} + 2V_{family(pop)})$,

Q_{CT} (among regions): $V_{region} / (V_{region} + V_{pop(region)} + 2V_{family(pop)})$

Q_{SC} (between populations within regions): $V_{pop(region)} / (V_{pop(region)} + 2V_{family(pop)})$

Pairwise Q_{ST}, for population pair i,j: $V_{i,j} / (V_{i,j} + 2V_{family(pop\ i,j)})$,

The variances V_{region}, V_{pop}, $V_{pop(region)}$ and $V_{family(pop)}$ were obtained from a nested linear model with three random factors: region (center, middle, edge), population nested within region, and family nested within population. Variances used to calculate Q_{ST} for each population pair (15 combinations) were estimated from a nested linear model with two factors: population and family nested within population.

For neutral quantitative traits that evolve primarily through drift, Q_{ST} is predicted to equal F_{ST} (Lande, 1992; Spitze, 1993; MacKay & Latta, 2002; Whitlock, 2008). For traits under the influence of diversifying (locally adaptive) selection among populations, Q_{ST} is predicted to be greater than F_{ST}. For traits under stabilizing selection for the same optimum across populations, Q_{ST} will be lower than F_{ST}. In our analyses, we compared each hierarchical Q statistic for each trait, to the average value across loci of each hierarchical F statistic. As a more rigorous test, we also compared the average pairwise Q_{ST} value for each trait to the distribution of all pairwise F_{ST} values for the eight loci. We considered traits with average pairwise Q_{ST} greater than the 95th percentile of the pairwise F_{ST} distribution to be under diversifying selection. We considered

average Q_{ST} between the 90th and 95th percentiles of the F_{ST} distribution as suggestive of diversifying selection (Whitlock, 2008).

Results

Landscape environmental variation. We found the environment varies from the western to the eastern range edge in a complex way. There exist two major axes of environmental variation in opposing geographic directions (Fig 2, Fig S2). Temperature and slope decrease in parallel (r = 0.56) from southwest to northeast across the study area (Fig. S2). Temperature declines because of increasing elevation, whereas the average slope of sites declines because the Kern Canyon is narrowest and steepest at its base in the southwest but broadens at the top and outside the main canyon. Spring precipitation and linear azimuth are correlated (r = 0.35) and vary in an orthogonal direction to the previous gradient, increasing from southeast to northwest across the range. Spring precipitation tends to be higher and sites tend to be more south facing in the northwestern section of the range. A gradient of solar radiation overlays variation in the previous variables, increasing strongly from west to east, and south to north due to the topography of the canyon (Figs. 2 and S2).

As a result of these complex and countervailing gradients, *Clarkia* populations experience a wide variety of environmental conditions across the range. At the extremes, some southwestern sites occur deep in the shaded parts of the canyon where precipitation is low, temperatures are high in the spring, and slopes are steep and north facing. Other populations occur east of the canyon where it is cooler, slopes are shallower and there is more sunlight.

Phenotypic trait differentiation between populations. For 30 populations measured both in the

field and greenhouse, all traits, except branchiness in the greenhouse, exhibited significant differentiation among populations (Table 1). The proportion of phenotypic variation attributable to population differences was highest for flowering time (47.4% field and 26.9% common garden) and lowest for branchiness (9.5% field and 0% common garden) (Table 1). We found the same result when examining six of these populations in a family-structured common garden with larger sample sizes of individuals within populations: All traits, this time including branchiness and two floral characters, exhibited significant differentiation among populations (Table 2). Flowering time, again, was the most differentiated trait. Phenotypic differentiation observed in the field persisted in both common gardens, indicating that populations are genetically differentiated for vegetative (height, branchiness, seed weight), phenological (flowering time), and reproductive (flower number, petal size, herkogamy) characters.

Nonetheless, there were clear effects of the growth environment on plant phenotypes. First, plants grown in the common garden were generally much larger, branchier, and had more flowers and heavier seeds than field plants (Table S1). Furthermore, population mean phenotypes in the field were not generally predictive of mean phenotype in the greenhouse, except for flowering time (Fig. 3, $P<0.0001$). Correlations between height, branchiness, and flower number between the field and greenhouse phenotypes were non-significant ($R^2=0.01$, 0.04, 0.11 respectively).

Environment-phenotype associations. All measured traits, except branchiness, co-varied with one or more environmental variables in the field (Table S2). The best environmental predictors of trait variation frequently were different in the field and greenhouse. In the field, phenotypic clines were most frequently related to variation in spring temperature (three traits) and/or linear

azimuth (two traits) among sites (Table S2, Fig. S3). By contrast, in the greenhouse, phenotypic clines were more often related to site differences in spring precipitation (three traits), solar radiation (three traits) and/or slope (two traits) (Table S2, Fig. S4). For example, in the field, plants were taller and had more flowers at warmer sites and/or sites with more south facing slopes (Fig. S3a, b, f), whereas in the greenhouse, height and flower number increased with spring precipitation and decreased with increasing solar radiation (Fig. S4a, b, g, i). The only trait with common environmental predictors in the field and common garden was flowering time. In both settings, flowering began earlier in populations from sites on steeper slopes with lower solar radiation (Fig. S3d & e vs. Fig. S4 e & f).

The fact that, for most traits, phenotypic clines are related to different environmental variables between field-grown and greenhouse-grown plants explains why height, branchiness and flower number in the field are poor predictors of these phenotypes in the greenhouse, and confirms the strong effect of growth environment on plant phenotypes. The exception to this is flowering time where population differences are correlated between the two growth environments (Fig. 4), and where clinal variation in both environments is related to the common environmental variables of slope and solar radiation (Table S2; Fig. S3d & e vs. Fig. S4e & f).

Trait heritabilities and genetic correlations between traits

In the family-structured common garden experiment, all traits, including two floral characters, exhibited significant heritable variation within one or more populations (Fig. 4). Indeed, within-population variation exceeded among-population variation for all traits except flowering time (49.2 % among populations vs. 20.6% among families within populations). Broad-sense heritability of traits (H^2) varied widely by population for all traits (Fig. 4).

Heritability was significantly greater than zero in 4 or more (out of 6) populations for all traits except flowering time, for which we could detect heritability significantly greater than zero in only 2 populations. We did not detect a relationship between the magnitude of genetic variation (or H^2) within populations and the geographic location [longitude (easting) and latitude (northing)] or census size of populations.

Although our statistical power was somewhat limited, we did not detect any strong genetic correlations between traits in the common garden. Only 19 out of 90 genetic correlations significantly differed from 0 (Table S3). Furthermore, for no trait pair was the correlation consistently strong or in the same direction in all populations, and no single population had strong correlations for all traits.

Quantitative versus neutral genetic differentiation: *F* and *Q* statistics

In general, average Q_{ST}'s (pairwise and overall across populations) were consistently higher than average F_{ST} across neutral loci (Table 3). Average pairwise Q_{ST} (column 2) was significantly greater than average F_{ST} for branchiness, flowering time, herkogamy, and seed weight. Differentiation was especially strong for the first three traits where Q_{ST} exceeded the 90th percentile (herkogamy, branchiness) and 95th percentile (flowering time) of the distribution of average F_{ST} (Figs. 5 and S5). When tested independently, average pairwise Q_{ST} for branchiness and herkogamy fell outside the distribution of pairwise F_{ST} values for six of eight sequenced loci, and flowering time fell outside the distribution of all eight loci (Fig. S6).

The relative degree of differentiation at different geographic scales was similar when measured using neutral loci and using quantitative traits. Neutral and trait-based differentiation between populations within regions (F_{SC} and Q_{SC}) were higher than differentiation between

regions (Q_{CT} and F_{CT}) (Table 3, column 4 vs. 5) for the four traits with the highest average pairwise Q_{ST}'s. This perhaps reflects the complexity of the environmental and selective landscape that varies as much within as between regions.

Discussion

The goals of this study were to quantify the level of phenotypic and genetic variation between and within populations that span the core geographic range of the California endemic plant *C. x.* ssp. *xantiana*, and to gauge the contribution of diversifying selection to phenotypic diversity by testing whether differences between populations are both heritable and likely to be driven by local adaptation. We in turn use this information to gain perspective on the likelihood of model-based predictions that (1) maladaptive gene flow or (2) lack of quantitative genetic variation at the range edge help prevent this species from expanding beyond its eastern range boundary which has no apparent physical barriers to dispersal. Generally we find that neither prediction is well supported.

Population differentiation and adaptation

In *C. xantiana* we found heritable trait differentiation between populations is ubiquitous and likely to be driven by local adaptation to climate variables. Among a sample of 30 populations spanning the narrow core of the species' geographic range we found strong evidence for phenotypic and genetic differentiation in vegetative, phenological, and reproductive characters. In a second common garden study focused on comparing central to peripheral populations, branchiness, two additional floral characters related to mating system, as well as seed weight were also differentiated among populations. Thus, in spite of asymmetric gene flow

between central and peripheral populations, and between populations within geographic regions (Moeller *et al.*, 2011), populations are genetically differentiated in ecologically important quantitative traits.

Several lines of evidence suggest that much of this phenotypic differentiation reflects adaptive divergence in response to varying environmental conditions. First, environmental factors that are known to affect plant growth, survival, and reproduction – such as spring temperature, precipitation, solar radiation, and site slope and aspect (Eckhart *et al.*, 2010, 2011; Kramer *et al.*, 2011) – vary strongly across the species' range, setting the stage for diversifying selection for different trait optima. Second, population average trait differences are correlated with variation in environmental factors. In the field and especially in a common garden, trait differentiation is related to multiple environmental factors that vary in opposing or intersecting gradients across the landscape. However, because the best environmental predictors of population differences in phenotype differed between the field and greenhouse-grown plants, it is difficult to say whether the pattern of phenotypic expression in the field (i.e. the plastic response of genotypes to field environmental differences), the genetically based differences in phenotype in the common greenhouse environment, or both reflect adaptation. In earlier field reciprocal transplant studies of central and peripheral *C. x.* ssp. *xantiana* populations, plasticity (large transplant site and year (wet vs. dry) effects on phenotype) was found in some of the same phenotypic traits (Eckhart, Geber, and McGuire,2004), as was evidence of local adaptation (Geber & Eckhart, 2005). Measurements of selection on phenotypic traits in reciprocally transplanted populations may help discern the relative importance of plasticity vs. intrinsic genotypic differences to adaptation.

In this study, additional evidence that trait divergence is adaptive comes from

comparisons of the level of differentiation at quantitative traits (Q_{ST}) vs. neutral loci (F_{ST}) in a sample of six central to peripheral populations. Q_{ST} was significantly higher than average F_{ST} for four out of six traits, and for flowering time, branchiness and herkogamy, range-wide Q_{ST} exceeded the 90th percentile of the distribution of pairwise F_{ST}'s. The fact that locally adaptive selection commonly promotes inter-population trait diversity, particularly in wild plants, has been demonstrated across studies (Linhart and Grant, 1996, McKay & Latta, 2002; Leinonen & Merilla, 2008), and here this conclusion is supported even at a relatively small spatial scale, across the core of an exceptionally narrow geographic species range.

The best evidence for strong genetic determination and common environmental drivers of adaptive divergence in *C. xantiana* is for flowering time. It is the trait that 1) is most differentiated among populations both in the field and greenhouse, 2) shows a strong correlation in expression between the two growth environments, and 3) is correlated with common environmental predictors – spring solar radiation and site slope – in the two growth environments. In annuals, early flowering is thought to be adaptive as a means of completing the life cycle before the onset of severe environmental stress, such as heat or drought (Levitt, 1980; Rice & Mack, 1991; Aronson *et al.*, 1992; Stanton, Roy, and Theide, 2000). Indeed, in the genus *Clarkia,* populations and species from hot and/or dry environments consistently flower earlier, when grown in a common environment, than those from cooler and/or wetter locations (Moore & Lewis, 1965; Vasek, 1971; Vasek & Sauer, 1971; Jonas & Geber, 1999; Eckhart, Geber, and McGuire, 2004; Dudley, Mazer, Galusky, 2007).

Q_{ST}/F_{ST} comparisons also suggest that population differences in herkogamy, a trait related to mating system, reflect adaptive divergence. Previous work has found that herkogamy is reduced in eastern peripheral populations relative to central populations (Moeller, 2006),

although petal (flower) size does not differ (Eckhart, Geber, and McGuire, 2004). Consistent with these earlier findings, Q_{ST}/F_{ST} is high for herkogamy but does not differ significantly from 1 for petal size. Small flowers and limited herkogamy are both generally associated with evolution of the selfing habit in plants (reviewed, Wyatt, 1988). They are characteristic of *Clarkia* selfing taxa (Lewis & Lewis, 1955) including the subspecies *C. x*. ssp. *parviflora* (Fausto, Eckhart, and Geber, 2001; Eckhart Geber, and McGuire, 2004). Plants with selfing flowers may be selected because they provide reproductive assurance in environments where mates or pollinators are scarce, but selfing may also be a pleiotropic side-product of the evolution of faster maturation in drought-prone habitats (Stebbins, 1950; Guerrant Jr., 1989).

Branchiness is a third trait for which Q_{ST} analysis suggests local adaptation. Branching is a complex trait that can be influenced by selection for light and nutrient use efficiency, or selection on root to shoot allocation ratio, among others (Bonser & Aarssen, 1996; Tiffin, 2000; Baker, Hileman, and Diggle, 2012). In our common garden study, branchiness was greater in populations from sites with higher precipitation. In years when water and other resources are abundant in the field (a condition simulated by the greenhouse environment), plants that produce more branches and hence more flowers, may have higher fitness than less branched plants. Average branchiness may also differ between populations due to selection caused by herbivory. *Clarkia xantiana* is eaten by several species of small mammals, and browsed, sometimes intensively, by mule deer and cattle (pers. obsv.). Among populations, from 5% to >50% of plants in the field were damaged by herbivores (data not shown). Within a population, the percentage of (ungrazed) plants with greater than zero branches increased with the site level of herbivory ($p=0.0005$, data not shown). Herbivory may select for plants with a high constitutive level of lateral meristem activity, which confers both general branchiness under normal

conditions and an increased compensatory growth response (Bonser & Aarssen, 1996; Strauss & Agrawal, 1999; Tiffin, 2000).

Other recent work has supported that *C. xantiana* populations are locally adapted, to varying extents. In several field reciprocal transplant studies (Geber & Eckhart, 2005), including the six populations from the family-structured common garden (Geber et al., in prep), central populations have consistently exhibited better performance in their home sites than at the range edge, eastern peripheral populations less so. With respect to fitness derived from immediate seed germination, survival, and reproduction, central populations appear to exhibit higher levels of local adaptation than peripheral populations, which only outperform central populations in edge transplant gardens in very dry years (Geber & Eckhart, 2005).

Could lack of genetic variation limit range expansion?

A potential evolutionary limit on range expansion is low levels of genetic variation in ecologically important traits at the range edge. If peripheral populations are demographically stable they are expected to diverge adaptively from central populations, but adaptation to environments beyond the range edge may be limited because a history of strong selection, small populations size and/or inbreeding has reduced levels of genetic variation in ecologically important traits. Similarly, adaptation may be limited by a lack of diversity in the way that traits are correlated with each other, that is if within a population there are strong maladaptive genetic correlations among traits. In *C. xantiana* more than one measure of overall genetic diversity (haplotype richness at nine nuclear loci and allelic richness at four microsatellite loci) are lowest in peripheral populations (and subsequently increase toward the range center) (Moeller *et al.*, 2011). However, in this study we found no evidence that the amount of genetic variation

(heritability) for traits or the number of traits with heritable variation was lower in peripheral compared to central populations. Variation was also not correlated with census population size. Most genetic correlations were not strong and the pattern of trait correlations was not consistent among populations, suggesting that the genetic architecture of traits is free to evolve. This result supports the long-standing argument that molecular diversity at putatively neutral loci does not provide a good estimate of adaptive diversity in many species, and should not be used exclusively in making conservation decisions (Reed & Frankham, 2001).

It is of note, however, that genetic constraints of other types may still contribute to lack of range expansion in *C. xantiana*. In particular, if many traits must be selected simultaneously to allow range expansion, lack of variation in some traits (particularly those not measured here) may still be low enough to prohibit colonization beyond the range edge despite high levels of variation in other required traits. Whether genetic constraints in general play a role in restricting adaptation varies broadly by species (Etterson & Shaw, 2001; Futuyma & Agrawal, 2009; Futuyma, 2010), and a role for genetic constraints cannot be ruled out for *C. xantiana*. Studies that have identified the genes that determine variation in locally adaptive traits and directly examined their patterns of variation, linkage, and flow between natural meta-populations are rare (but see Hohenlohe *et al.*, 2010; Hanski, 2011).

Could maladaptive gene flow limit range expansion?

Source-sink or colonization/extinction dynamics also have the potential to limit adaptation if populations at the range edge persist solely through immigration from central ones. Edge populations may be perpetually maladapted and genetically similar to more central populations because of repeated or swamping gene flow from the range center. In *C. xantiana,*

the results of this study are not definitive, but certainly do not strongly support this hypothesis. Despite gene flow between populations, edge populations are quite differentiated from more central ones and the amount of heritable trait variation in edge populations is the same, or in some cases greater than, variation in more central populations (Fig. 4). The fact remains however, that we cannot explicitly rule out that gene flow has any role at all in inhibiting range edge adaptation using current theoretical models. The most well-known models address the dynamics of adaptation across linear environmental gradients (Lande, 1992; Dias, 1996; Kirkpatrick & Barton, 1997) but have not, to our knowledge, addressed adaptation across more complex landscapes or selection scenarios. For example, an assumption of the Kirkpatrick and Barton (1997) model is that trait optimums across populations follow a linear gradient. In the present study we find that the most differentiated trait, flowering time, varies significantly along two or more environmental axes in both the greenhouse and the field and these axes are themselves not parallel but sometimes orthogonal. The model also requires that selection on traits be stabilizing in all populations. Work involving reciprocal transplants of the same six *C. xantiana* populations from this study (unpublished) shows that selection on flowering time is not stabilizing at all sites, nor does the optimum (if one exists) follow a simple linear environmental gradient.

It is true that across simple linear environmental gradients, a low amount of gene flow from range center to edge may allow population differentiation but still prevent trait variation in range edge populations from attaining the right magnitude to allow adaptation to conditions beyond the edge (Dias, 1996; Kirkpatrick & Barton, 1997). The same might be predicted for plants distributed across landscapes encompassing more than one significant environmental gradient, as in this study. We might reasonably expect the genetic trait variation and covariation

required for adaptation in such cases is much greater than that required for adaptation at the extreme of a single gradient. Few studies have directly tested whether the magnitude or rate of change of phenotypic variation from center to edge populations excludes the possibility of adaptation at the range edge (but see Angert & Schemske, 2008). However, because for *C. xantiana* the amount of heritable trait variation in edge populations is within the same range of values and varies in some cases more than in central populations, we find it unlikely that adaptation would be limited because of maladaptive gene flow in *C. xantiana*.

In the present day, increasingly researchers are working towards direct characterization of the relationship between genetic variation and adaptive phenotypic differences by associating genotypes with phenotypes using high-throughput genotyping methods (Stinchcombe & Hoekstra, 2008; Nadeau & Jiggins, 2010; Stapley *et al.*, 2010). Directly measuring the frequencies of functionally adaptive alleles in range center and edge populations may be a promising area of research in the continued analysis of the role of gene flow and local adaptation in range boundary formation in wild plant species.

Acknowledgements

We thank Q. Liang and J. Hillabrant for help in the field. We thank A. Erwin, S. Cook-Patterson, M. Weber, I. Singh, and R. Petipas for insightful comments on the manuscript. The research was supported by National Science Foundation grant DEB-0515428 to M.A.G., DEB-0515466 to P.T. and D.A.M., and DEB-0515409 to V.M.E.

FIGURES

Figure 1. a) Location of 30 populations (circles) surveyed in the field near Lake Isabella, CA. Elevation in the region ranges from <500 m to >2500 m, indicated by increasing shading in 500 m intervals. Filled (white) symbols represent the six populations used in the family-structured common garden study. b) Morphological variation among similar aged plants grown in a common garden. Inset: *C. xantiana* flower.

(a) (b)

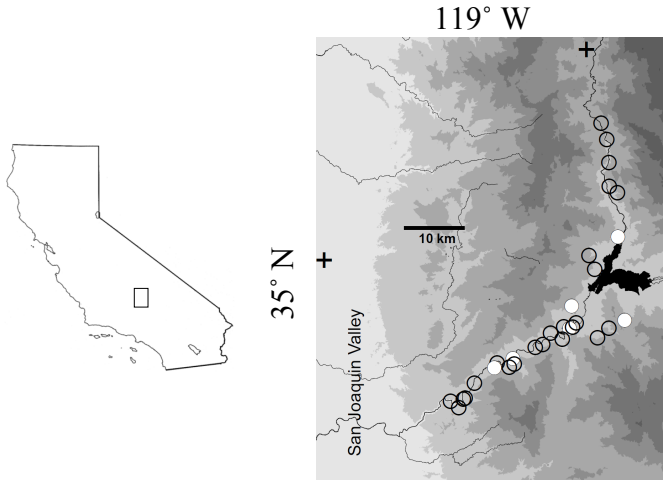
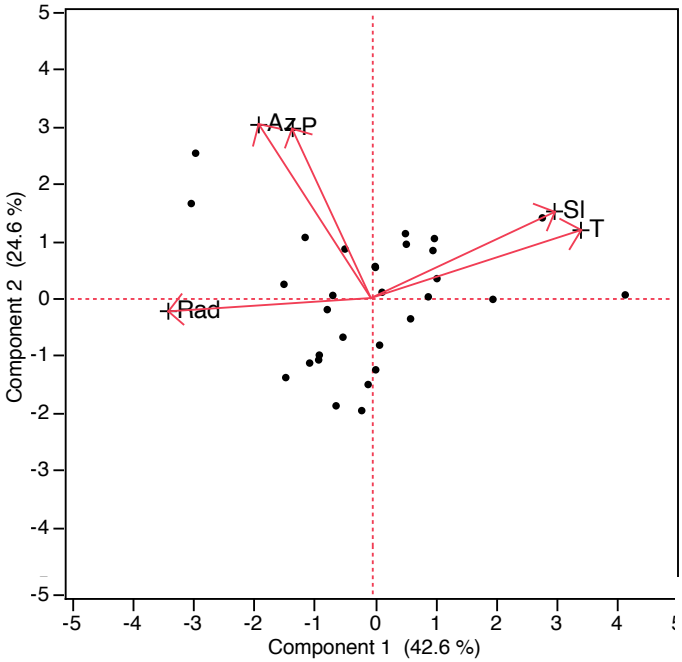


Figure 2. PCA ordination plot of environmental conditions: Az, linear azimuth; P, spring precipitation; T, temperature; Sl, slope; Rad, solar radiation. Table shows variable loadings on principle components and component eigenvalues.



	Prin1	**Prin2**	Prin3	Prin4	Prin5
slope (deg inclination)	**0.723**	0.358	0.474	0.088	-0.339
linear azimuth (deg from N)	-0.441	**0.722**	0.423	-0.231	0.222
solar radiation (WH/sq m)	**-0.796**	-0.055	0.301	0.521	-0.019
spring temperature (deg C)	**0.827**	0.282	-0.124	0.337	0.326
spring precipiation (cm)	-0.309	**0.703**	-0.603	0.098	-0.188
Eigenvalue	2.132	1.2288	0.875	0.456	0.307
Variation explained	42.6%	24.6%	17.5%	9.1%	6.1%
Prob>ChiSq	0.003	0.062	0.143	0.568	n/a

Figure 3. Population average flowering time in the common garden and the field for 30 *Clarkia* populations.* Index of flowering in the field (flowering stage) has been multiplied by negative 1.

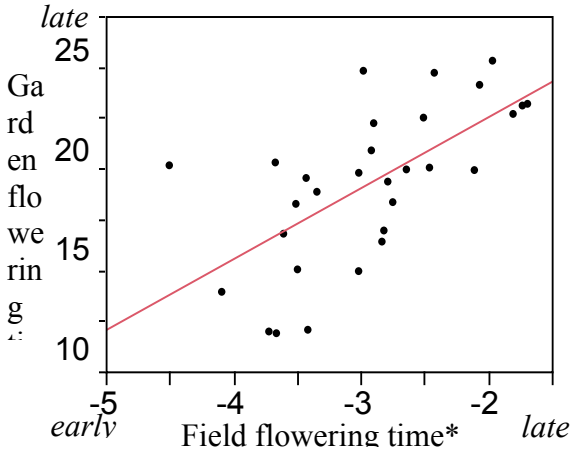


Figure 4. Broad-sense heritability estimates for six plant traits for six populations. Error bars represent +/- one standard error. Populations are ordered from southwest to northeast (range center to range edge) for each trait. *Value significantly different than zero.

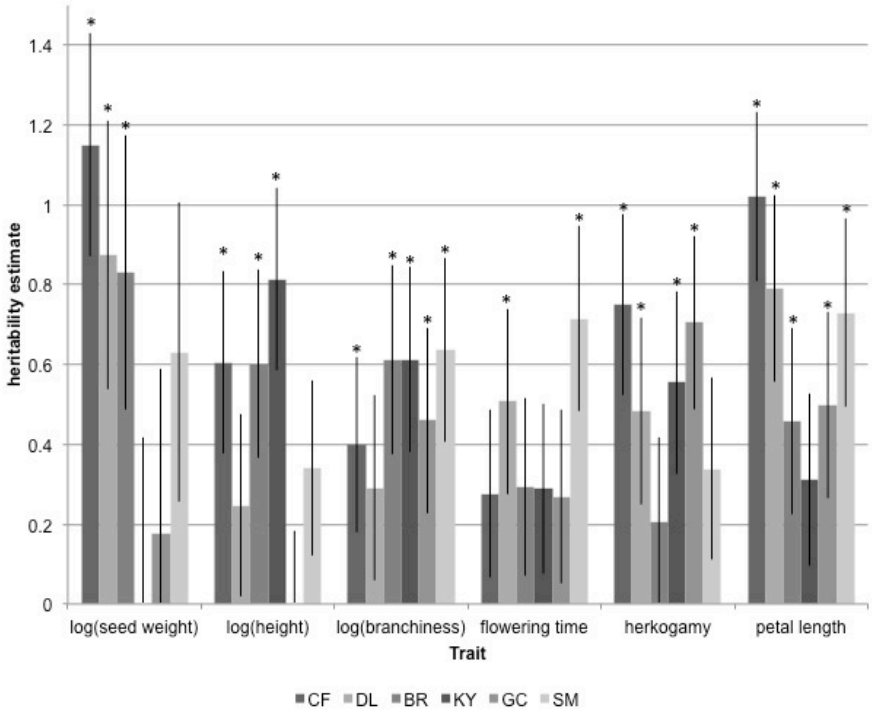
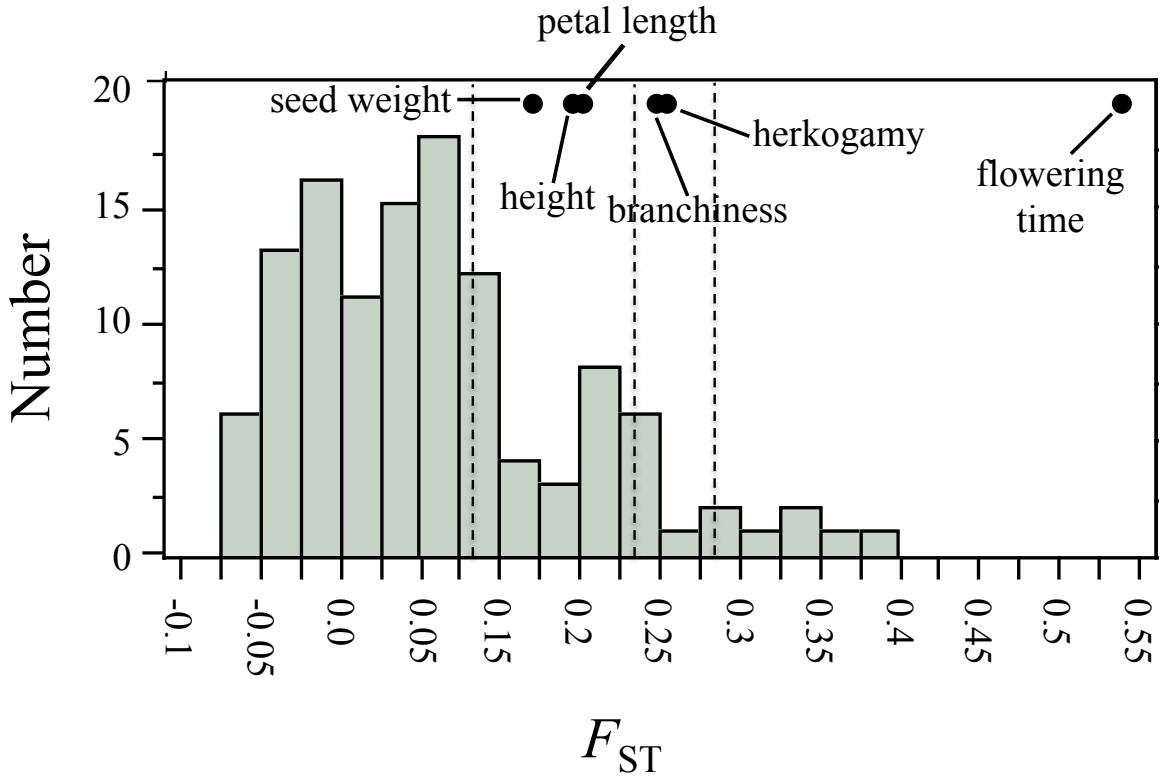


Figure 5. Frequency distribution of pairwise F_{ST} values for eight molecular markers in six populations of *Clarkia* used in the family-structured common garden study (N=121 values, see Methods). The dotted lines mark the mean, the 90th, and the 95th percentiles of the distribution (0.109, 0.235 and 0.285, respectively). Dots represent the average pairwise Q_{ST} values for each trait. See Fig. S5-6 for separate F_{ST} or Q_{ST} distributions for each locus and trait.



TABLES

Table 1. Field and common garden variance components of phenotypic traits for 30 populations.

	Field 2008			Common Garden 2008		
Phenotypic Trait	Total [a]	B/w pop (%) [b]	W/in pop [c]	Total [a]	B/w pop (%) [b]	W/in pop [c]
Height (cm) (log_e)	0.258	**0.056** (21.6)	0.202	0.076	**0.016** (20.5)	0.060
Branchiness (no./cm) (log_e)	0.345	**0.036** (9.5)	0.345	0.101	0 (0)	0.101
Flowering time [d]	1.083	**0.512** (47.4)	0.570	44.38	**11.94** (26.9)	32.44
Flower number (log_e)	0.861	**0.105** (12.2)	0.755	0.389	**0.056** (14.5)	0.333
Seed weight (mg) (log_e)	0.066	**0.025** (37.4)	0.041	--	--	--

[a] The total phenotypic variance is the sum of the variance within and between populations.
[b] Phenotypic variance (and percentage of total) attributable to differences between populations. Significant values in bold ($p=<0.05$).
[c] Phenotypic variance within populations.
[d] Measured as adjusted flowering stage in the field, as date of first flowering in the common garden.

Table 2. Components of phenotypic variation among six populations in a family-structured common garden.

	Total Phenotypic Variation (V_P)			
Phenotypic Trait	Total [a]	B/w pop (%) [b]	B/w families (%) [c]	W/in families
Height (cm) (log_e)	0.050	**0.004** (8.1)	**0.009** (20.1)	0.037
Branchiness (no./cm) (log_e)	0.265	**0.049** (18.3)	**0.053** (24.8)	0.163
Flowering time [d]	27.32	**13.46** (49.2)	**2.86** (20.2)	11.00
Seed weight (mg) (log_e)	0.041	**0.004** (10.5)	**0.011** (31.3)	0.025
Anther-stigma distance (mm)	0.051	**0.009** (17.4)	**0.010** (24.9)	0.031
Petal length (cm)	0.084	**0.013** (15.2)	**0.022** (30.5)	0.049

[a] Total phenotypic variance is the sum of the variances between populations, between families within populations, and within families.
[b] Phenotypic variance (and percentage of total) attributable to differences between populations. Significant values in bold ($p=<0.05$).
[c] Phenotypic variance between families within populations, and in parenthesis the percentage of variance between families as a fraction of the sum of the variance within and between families. Significant values in bold ($p=<0.05$).
[d] Measured as date of first flowering in the common garden.

Table 3. Pairwise and hierarchical Q and F statistics among six greenhouse populations.

Trait	Average pairwise Q_{ST} and F_{ST}	Between populations Q_{ST} and F_{ST}	Between populations within regions Q_{SC} and F_{SC}	Between regions Q_{CT} and F_{CT}
Height ($\log_e$)	0.187	0.18	0.182	0
Branchiness ($\log_e$)	0.246	0.312	0.301	0
Flowering time	0.532	0.697	0.69	0.084
Seed weight ($\log_e$)	0.171	0.159	0.058	0.13
Herkogamy	0.256	0.297	0.291	0
Petal length	0.192	0.229	0.078	0.192
F_{ST}	0.109	0.082	0.049	0.034

SUPPLEMENTAL FIGURES and TABLES

Figure S1. Contour plots of environmental variables across the core geographic range. Plot shading represents equal probability contours interpolated from environmental data using 2-dimensional density estimation. Dots are the values/locations of the 30 sampled populations.

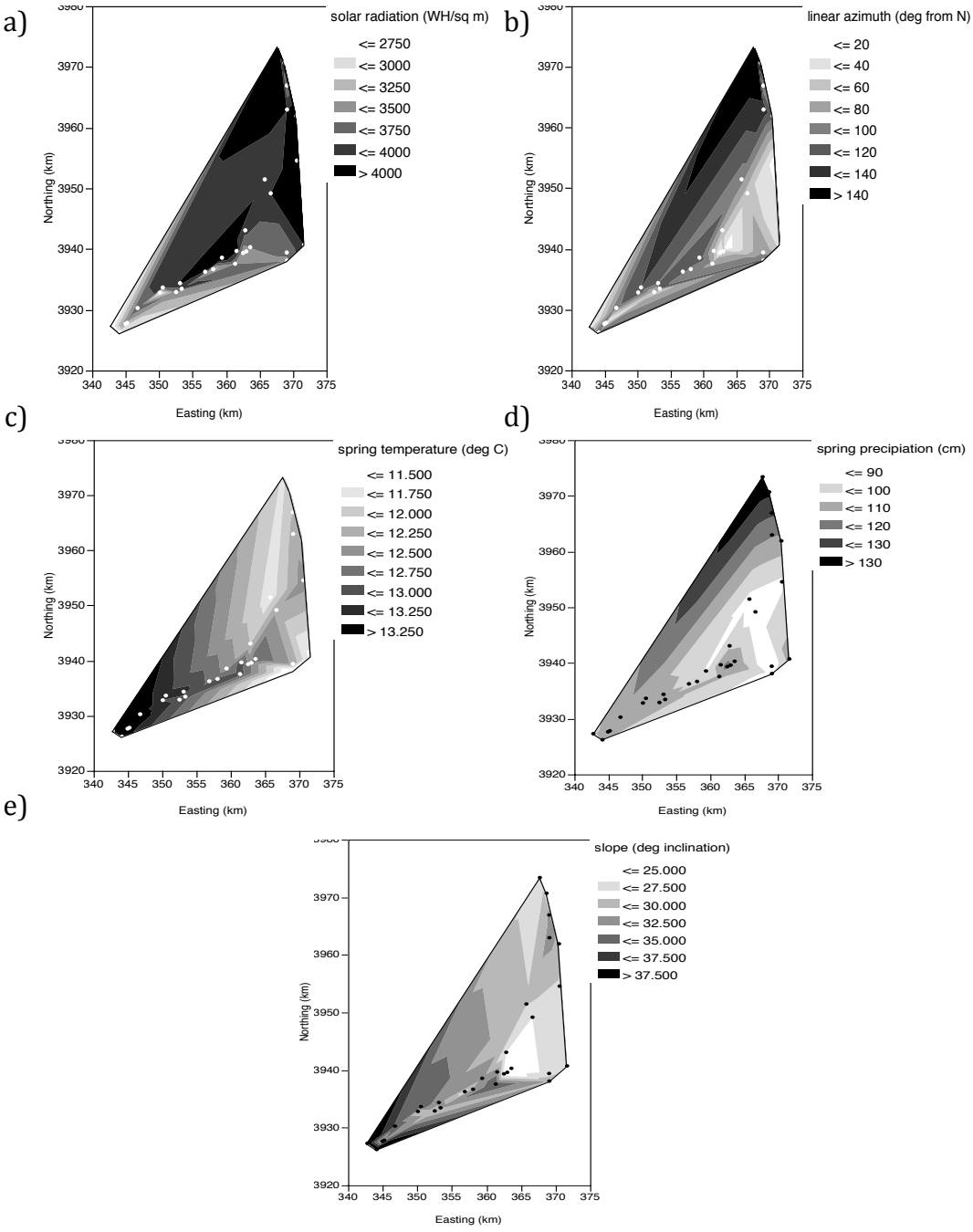


Figure S2. Regression and partial regression plots of average population plant trait values in the field versus field environmental variables. Each plots represents the relationship between a plant trait and an environmental variable, holding all other environmental variables constant (see methods). *Larger residual flowering stage indicates *earlier* flowering.

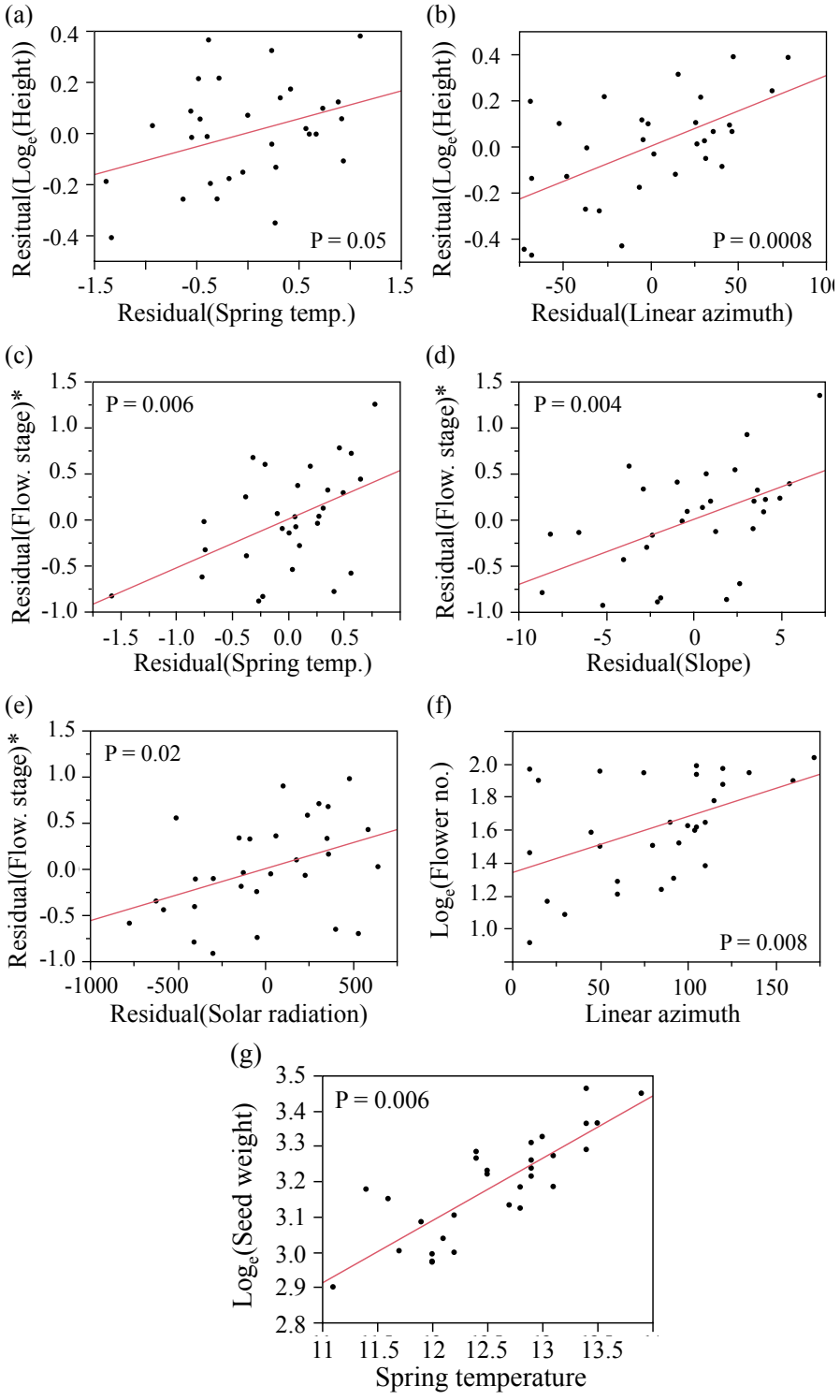


Figure S3. Partial correlation plots of average population plant trait values in the common garden versus field environmental variables in *Clarkia*. Each plot represents the relationship between a plant trait and a single environmental variable, holding all other environmental variables constant (see methods). *Larger residual flowering time indicate *later* flowering.

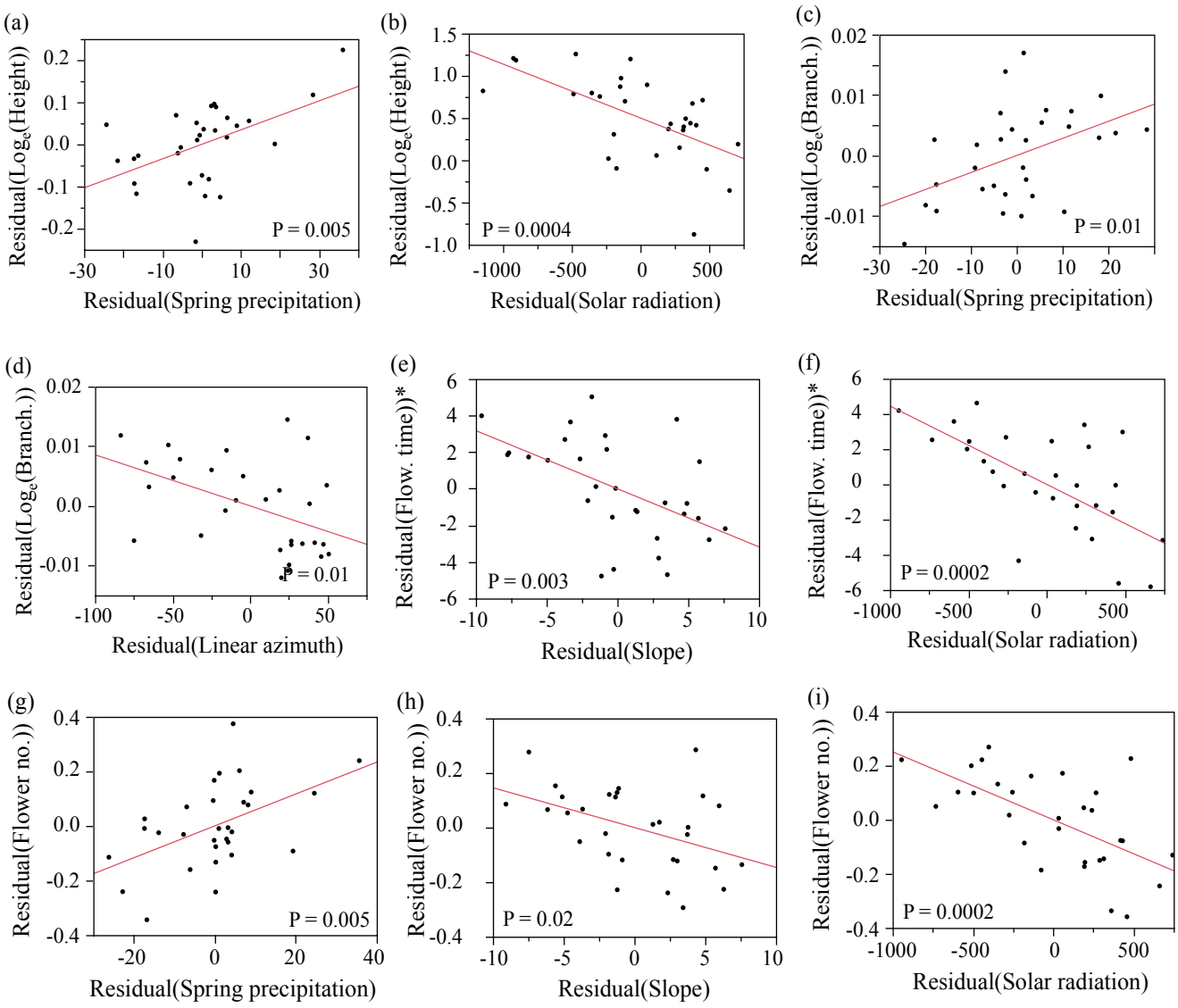


Figure S4. Pairwise F_{ST} and Q_{ST} distributions. Light grey bars represent the distribution of population pairwise F_{ST} values (n =121) and dark grey bars represent the distribution of pairwise Q_{ST} values for an individual trait (n = 14-15 for each trait). Dotted lines represent the average of Q_{ST} values for each trait.

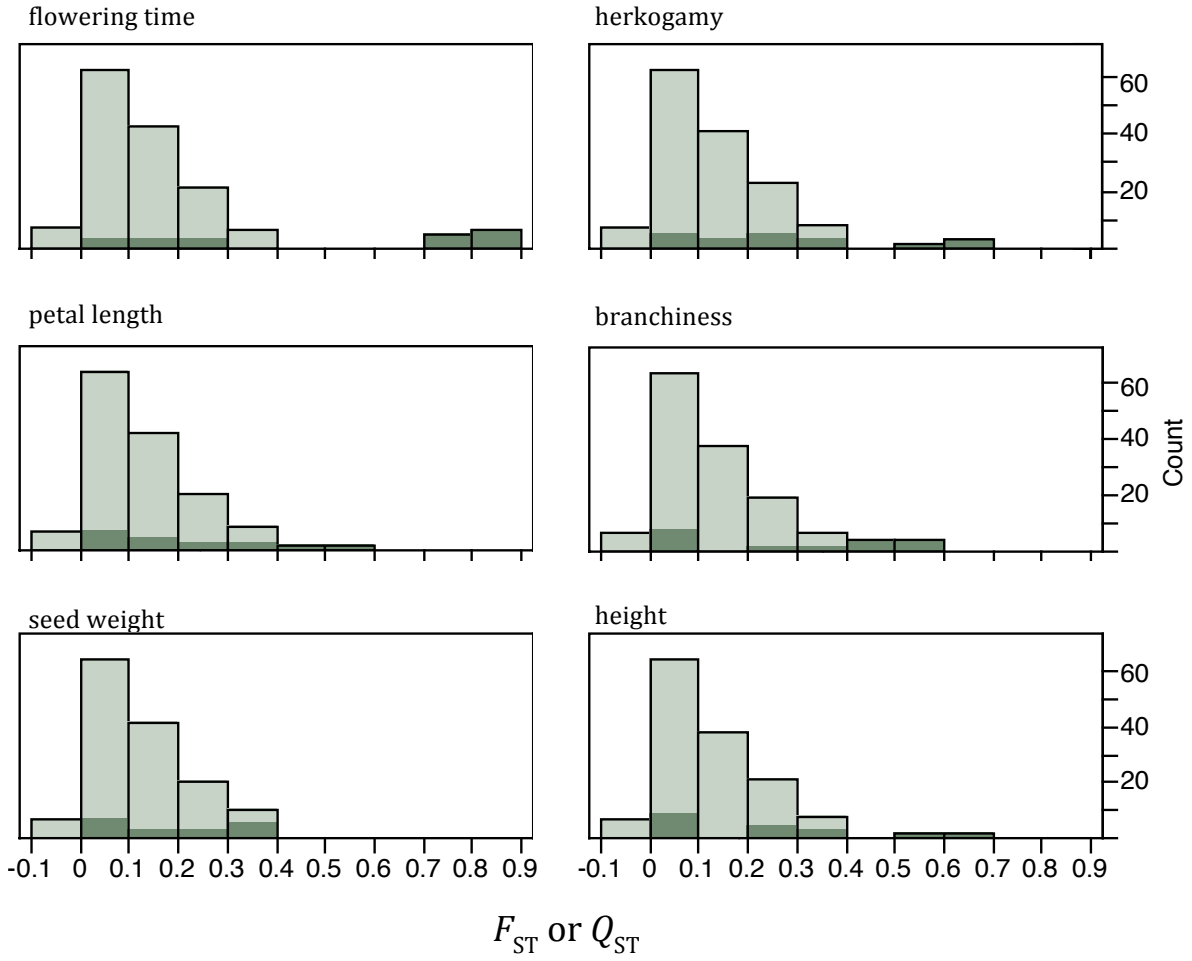


Figure S5. Pairwise F_{ST} distributions for each of 8 population genetic markers. Dotted lines show the mean pairwise Q_{ST} for each trait: ht, height; sw, seed weight; pl, petal length; br, branchiness; ft, flowering time.

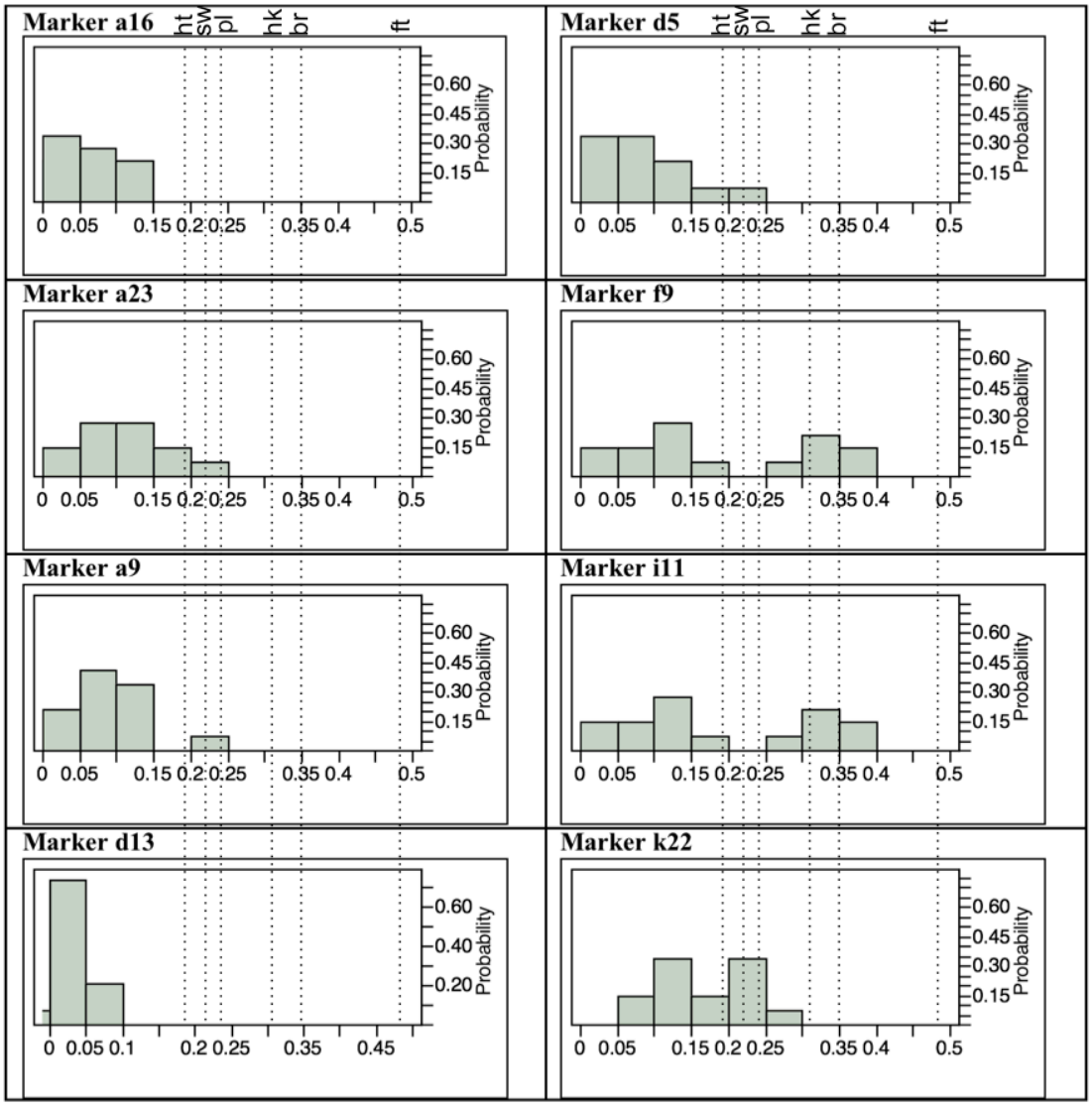


Table S1. Ranges of population average trait values.

Phenotypic Trait	Field (N=30)	Common Garden (N=30)
Height (cm)	19 - 46	51 - 79
Branchiness (br no./cm)	0.035 - 0.061	0.129 - 0.125
Time to flowering (stage or days)	1.69 - 4.50 (approx. 0-28 days)	0 - 22
Flower number	2.5 - 7.7	26.2 - 55.1
Seed weight (mg)	0.018 – 0.032	0.040 – 0.046 (N=6)[a]

[a] Data from the six populations in the family-structured common garden.

Table S2. Environmental variables of significant effect on plant traits in stepwise regressions.

		Field		**Common Garden**	
Phenotypic trait	Environmental variable	Estimate ($x10^2$)	Prob>F	Estimate ($x10^2$)	Prob>F
Height (log_e)	Spring temperature	10.9	**0.05**	-	-
	Spring precipitation	-	-	0.34	**0.006**
	Linear azimuth	0.31	**0.001**	-	-
	Solar radiation	-	-	-0.01	**0.001**
Branchiness (log_e)	Spring precipitation	-	-	0.03	**0.011**
	Linear azimuth	-	-	0.009	**0.011**
Flower stage/time	Spring temperature	52.81	**0.008**	-	-
	Slope	7.06	**0.005**	-31.79	**0.003**
	Solar radiation	0.06	**0.027**	-0.44	**0.000**
Flower no. (log_e)	Spring precipitation	-	-	0.58	**0.007**
	Linear azimuth	0.34	**0.008**	-	-
	Slope	-	-	-1.46	**0.020**
	Solar radiation	-	-	-0.03	**0.000**
Seed weight (log_e)	Spring temperature	17.67	**0.000**	N/A	

Table S3. Genetic correlations between pairs of traits within six populations. Values in bold are significant at $p<0.05$, values in italic are significant at $p<0.10$.

Trait 1	Trait 2	Population					
		CF	DL	BR	KY	GC	SM
Mean(log(branchiness))	Mean(log(height))	-0.07	0.16	**-0.54**	**-0.57**	0.13	*-0.36*
Mean(days to flowering)	Mean(log(height))	**-0.65**	-0.34	**-0.52**	**-0.55**	**-0.44**	-0.06
Mean(days to flowering)	Mean(log(branchiness))	0.30	0.19	*0.35*	*0.36*	**0.48**	0.17
Mean(herkogamy)	Mean(log(height))	**0.53**	-0.17	0.17	-0.06	**-0.52**	0.01
Mean(herkogamy)	Mean(log(branchiness))	-0.11	0.15	0.10	-0.09	0.23	0.27
Mean(herkogamy)	Mean(days to flowering)	-0.27	**0.48**	-0.17	*0.39*	**0.64**	-0.07
Mean(petal_size)	Mean(log(height))	**0.47**	0.16	**0.51**	-0.02	0.13	*0.36*
Mean(petal_size)	Mean(log(branchiness))	*-0.38*	-0.13	-0.33	-0.11	-0.11	-0.30
Mean(petal_size)	Mean(days to flowering)	**-0.49**	**-0.53**	-0.12	-0.32	**-0.53**	**-0.42**
Mean(petal_size)	Mean(herkogamy)	**0.52**	*-0.40*	0.10	-0.10	-0.16	0.30
Mean(Res_log_seed_weight)	Mean(log(height))	0.32	0.30	0.12	0.25	-0.21	-0.04
Mean(Res_log_seed_weight)	Mean(log(branchiness))	0.03	0.21	-0.15	-0.17	-0.17	-0.14
Mean(Res_log_seed_weight)	Mean(days to flowering)	-0.24	0.08	0.17	0.18	0.00	*0.36*
Mean(Res_log_seed_weight)	Mean(herkogamy)	0.23	0.05	-0.06	**0.41**	0.24	-0.04
Mean(Res_log_seed_weight)	Mean(petal_size)ß	0.06	-0.07	0.15	*-0.35*	-0.22	0.14

Literatire Cited

Angert A.L., Schemske D.W. (2008) Using experimental evolution to investigate geogrphic range limits in monkeyflowers. *Evolution* **62**, 2660–2675.

Aronson J, Kigel J, Shmida A, Klein J. (1992) Adaptive phenology of desert and Mediterranean populations of annual plants grown with and without water stress. *Oecologia,* **89**, 17–26.

Baker R.L., Hileman L.C., Diggle P.K. (2012) Patterns of shoot architecture in locally adapted populations are linked to intraspecific differences in gene regulation. *New Phytologist*, **196**, 271–281.

Bonser S.P, Aarssen L.W. (1996) Meristem allocation : A new classification theory for adaptive strategies in herbaceous plants. *Oikos,* **77**: 347–352.

Bridle J.R., Vines T.H. (2007). Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology & Evolution,* **22**, 140–147.

Brown J.H., Lomolino M.V. (1998) *Biogeography*. Sinauer, Sunderland, MA.

Dias, P. (1996). Sources and sinks in population biology. *Trends in Ecology and Evolution,* **8**, 326-330.

Dudley L.S., Mazer S.J., Galusky P. (2007) The joint evolution of mating system, floral traits and life history in *Clarkia* (Onagraceae): genetic constraints vs. independent evolution. *Journal of Evolutionary Biology*, **20**, 2200–18.

Eckert C.G., Samis K.E., Lougheed S.C. (2008) Genetic variation across species at geographical ranges: the central-marginal hypothesis and beyond. *Molecular Ecology,* **17**, 1170–1188.

Eckhart V.M., Geber M.A. (1999) Character variation and geographic distribution of *Clarkia xantiana* A. Gray (Onagraceae): flowers and phenology distinguish two subspecies. *Madroño,* **46**, 117–125.

Eckhart V.M., Geber M.A., McGuire C.M. (2004) Experimental studies of adaptation in *Clarkia xantiana*. I. Sources of trait variation across a subspecies border. *Evolution,* **58**, 59–70.

Eckhart V.M., Geber M.A., Morris W.F., Fabio E.S., Tiffin P., Moeller D.A. (2011) The geography of demography: Long-term demographic studies and species distribution models reveal a species border limited by adaptation. *American Naturalist,* **178**, S26–S43.

Eckhart V.M., Singh I., Louthan A., Keledjian A., Chu A., Moeller D., Geber M. (2010.) Plant-soil water relations and species border of *Clarkia xantiana* ssp. *xantiana* (Onagraceae). *International Journal of Plant Sciences,* **171**, 749–760.

Etterson J.R., Shaw R.G. (2001) Constraint to adaptive evolution in response to global warming. *Science,* **294**, 151–154.

Excoffier L., Smouse P.E., Quattro J.M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics,* **131**, 479–491.

Falconer DS. (1960). *Introduction to Quantitative Genetics*. New York: The Ronald Press.

Fausto JA, Eckhart VM, Geber MA. (2001) Reproductive assurance and the evolutionary ecology of self-pollination in *Clarkia xantiana* (Onagraceae). *American Journal of Botany,* **88**, 1794–1800.

Futuyma D.J. (2010) Evolutionary constraint and ecological consequences. *Evolution* **64**: 1865–1884.

Futuyma D.J, Agrawal A.A. (2009) Macroevolution and the biological diversity of plants and herbivores. *Proceedings of the National Academy of Sciences,* **106**, 18054–18061.

Gaston K.J. (2003) *The structure and dynamics of geographic ranges.* Oxford University Press, Oxford.

Geber M.A. (2011) Ecological and evolutionary limits to species geographic ranges. *American naturalist,* **178**, S1–5.

Geber MA, Eckhart VM. (2005) Experimental studies of adaptation in *Clarkia xantiana*: II. Fitness variation across a subspecies border. *Evolution* **59**, 521–531.

Grinnell J. (1917) The niche-relationships of the California Thrasher. *Auk* **34**, 427–433.

Guerrant Jr E.O. (1989) Early maturity, small flowers, and autogamy; A developmental connection? *The Evolutionary Ecology of Plants* (eds J.H. Bock & Y.B. Linhart), pp. 61–84. Westview,, Boulder, CO.

Hanski I.A. (2011) Eco-evolutionary spatial dynamics in the Glanville fritillary butterfly. *Proceedings of the National Academy of Sciences,* **108**, 14397–14404.

Hoffmann AA, Blows MW. (1994) Species borders: ecological and evolutionary perspectives. *Trends in Ecology & Evolution* **9**, 223–227.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA. (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet* **6**, e1000862.

Hudson RR, Boos DD, Kaplan NL. (1992) A statistical test for detecting geographic subdivision. *Molecular Biology and Evolution.* **9**, 138–51.

Hurvich CM, Tsai C-L. (1989) Regression and time series model selection in small samples. *Biometrika* **76**, 297.

Jonas CS, Geber MA. (1999) Variation among populations of *Clarkia unguiculata* (Onagraceae) along altitudinal and latitudinal gradients. *American Journal of Botany.* **86**, 333–343.

Keller, S, Soolanayakanahally, R, Guy, RD, Silim, SN, Olson, MS, Tiffin, P. (2011) Climate-driven local adaptation of ecophysiology and phenology in balsam poplar, *Populus balsamifera* L. (Salicaceae). *American Journal of Botany.* **1**,99-108.

Kirkpatrick M, Barton NH. (1997) Evolution of a Species' Range. *American Naturalist* **150**, 1–23.

Kramer HAC, Montgomery DM, Eckhart VM, Geber MA. (2011) Environmental and dispersal controls of an annual plant's distribution, how similar are patterns and apparent processes at two spatial scales? *Plant Ecology* **212**, 1887–1899.

Lande R. (1992) Neutral Theory of Quantitative Genetic Variance in an Island Model with Local Extinction and Colonization. *Evolution* **46**, 381–389.

Leinonen T, Merilla J. (2008) Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology* **21**, 1–17.

Levitt J. (1980) *Responses of plants to environmental stress*. Academic Press, New York.

Lewis H, Lewis ME. (1955) *The genus* Clarkia. University of California Press, Berkeley, CA.

Linhart, YB, Grant MC. (1996) Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecology and Systematics*. **27**, 237-277.

MacArthur RH. (1972) *Geographical ecology: patterns in the distribution of species*. Harper & Row, New York, NY.

Mayr E. (1963) *Animal species and evolution*. Harvard University Press, Cambridge, MA.

Mazer SJ, Paz H, Bell MD. (2004) Life history, floral development, and mating system in *Clarkia xantiana* (Onagraceae): Do floral and whole-plant rates of development evolve independently? *American Journal of Botany,* **91**, 2041–2050.

McKay JK, Latta RG. (2002) Adaptive population divergence: markers, QTL and traits. *Trends in Ecology and Evoltuion* **17**, 285–291.

Moeller DA. (2006) Geographic structure of pollinator communities, reproductive assurance, and the evolution of self-pollination. *Ecology,* **87**, 1510–1522.

Moeller DA, Geber MA, Eckhart VM, Tiffin PL. (2012) Reduced pollinator service and elevated pollen limitation at the geographic range limit of an annual plant. *Ecology,* **93**, 1036–48.

Moeller DA, Geber MA. (2005) Ecological context of the evolution of self-pollination in *Clarkia xantiana*: population size, plant communities, and reproductive assurance, *Evolution* **59**, 786–799.

Moeller DA, Geber MA, Tiffin P. (2011) Population genetics and the evolution of geographic range limits in an annual plant. *American Naturalist.* **178**, S44–S61.

Moore DM, Lewis H. (1965) The evolution of self-pollination in *Clarkia xantiana*. *Evolution* **19**, 104–114.

Nadeau NJ, Jiggins CD. (2010) A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. *Trends in genetics,* **26**, 484–92.

Pettengill JB, Moeller DA. (2012a) Tempo and mode of mating system evolution between incipient *Clarkia* species. *Evolution.* **66**, 1210–25.

Pettengill JB, Moeller DA. (2012b) Phylogeography of speciation: allopatric divergence and secondary contact between outcrossing and selfing *Clarkia*. *Molecular Ecology*. 21: 4578–4592.

Platt A, Horton M, Huang YS, Li Y, Anastasio AE, Mulyati NW, Agren J, Bossdorf O, Byers D, Donohue K, *et al. (*2010) The scale of population structure in *Arabidopsis thaliana*. *PLoS Genetics.* **2**,e1000843.

Pulliam HR. (1988) Sources, sinks, and population regulation. *American Naturalist* **132**, 652–661.

Reed DH, Frankham R. (2001) How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution.* **55**, 1095–103.

Rice KJ, Mack RN. (1991) Ecological genetics of *Bromus tectorum* III . The demography of reciprocally sown populations. *Oecologia*, 91–101.

Roff, Derek A, ed. (1997) *Evolutionary Quantitative Genetics*. Chapman and Hall., New York, NY.

Runions CJ, Geber MA. (2000) Evolution of the self-pollinating flower in *Clarkia xantiana* (Onagraceae). I. Size and development of floral organs. *American Journal of Botany,* **87**, 1439–1451.

Spitze K. (1993) Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics* **135**, 367–374.

Stanton ML, Roy BA, Thiede DA. (2000) Evolution in stressful environments. I. Phenotypic variability, phenotypic selection, and response to selection in five distinct environmental stresses. *Evolution.* **54**, 93–111.

Stapley J, Reger J, Feulner PGD, Smadja C, Galindo J, Ekblom R, Bennison C, Ball AD, Beckerman AP, Slate, J. (2010) Adaptation genomics: the next generation. *Trends in Ecology & Evolution* **25**, 705–12.

Stinchcombe JR, Hoekstra HE. (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* **100**, 158–70.

Stebbins GL. (1950) *Variation and Evolution in Plants*. Columbia University Press, New York.

Strauss S, Agrawal A. (1999) The ecology and evolution of plant tolerance to herbivory. *Trends in Ecology & Evolution* **14**, 179–185.

Tiffin P. (2000) Mechanisms of tolerance to herbivore damage: what do we know? *Evolutionary Ecology*, 523–536.

Vasek FC. (1971) Variation in marginal populations of *Clarkia*. *Ecology.* **52**, 1046–1051.

Vasek FC, Sauer RH. (1971) Seasonal progression of flowering in *Clarkia*. *Ecology* **52**, 1038–1045.

Via S. (1984) The quantitative genetics of polyphagy in an insect herbivore. II. Genetic correlations in larval performance within and among host plants. *Evolution.* **38**, 896–905.

Warren RJ. (2008) Mechanisms driving understory evergreen herb distributions across slope aspects: as derived from landscape position. *Plant Ecology* **198**, 297–308.

Whitlock MC. (2008) Evolutionary inference from Q_{ST}. *Molecular Ecology* **17**, 1885–96.

Wright S. (1951) The genetic structure of populations. *Annals of Eugenics*, 323–354.

Wyatt R. (1988) Phylogenetic aspects of the evolution of self-pollination. *Plant Evolutionary Biology*. (eds L. Gottlieb, D. Leslie, and K. Subodh), Chapman and Hall, London, UK.

CHAPTER 2

Development of cultivation and assay protocols for studying soil aluminum tolerance in the wild grass, *Anthoxanthum odoratum*

Abstract

Here I present the development of protocols for cultivating and testing the wild grass *Anthoxanthum odoratum* for tolerance to soil aluminum. I found that the species is easy to cultivate in the greenhouse and growth chamber both as transplanted tillers and as seedlings. Al tolerance is best measured in an aerated low ionic strength nutrient solution modified from the Long-Ashton solution used by previous researchers. Seedlings can be germinated and tested on this solution solidified with 1% agarose. Al tolerance can be measured at both life stages with multiple levels of experimental control by measuring relative root growth over 4 days each in control and Al exposure treatments. I found that this grass far surpasses other cultivated grasses for Al tolerance. It is best tested at least 300 uM activity Al^{3+} in solution. Using these methods this species can now be used to explore the ecological genetic basis of adaptation to soil Al in natural settings.

Introduction

The presence of mononuclear aluminum ions (Al^{3+}) in soil is a well known challenge faced by cultivated plants throughout the world, but the ways which wild plants have evolved to naturally deal with this challenge is little understood. Naturally occurring aluminum, the third most abundant mineral in the earth's crust, becomes phytotoxic in acidic soils (below pH 5.5), and acid soils are common, particularly in Asia and the Americas (Rengel, 2003). Aluminum toxicity has placed major constraints on cereal grass productivity and improvement of Al tolerance is a continuing goal for crop breeders. Al tolerance is an equally important trait in wild grasses that colonize acid soils and form the basis of both permanent and successional grassland ecosystems and rangelands (Rao *et al.*, 1993; Poozesh *et al.*, 2007). The phenotypic level of Al tolerance among wild grasses has been measured in some wild species (Foy, 1988; Wheeler *et al.*, 1992a; Rao *et al.*, 1993; Wheeler, 1995; Wenzl *et al.*, 2001; Poozesh *et al.*, 2007; Haling

et al., 2010), but the genetic basis of this critical adaptation remains entirely undissected at the genetic level in wild species, despite a wealth of useful comparative information in closely related cultivated ones.

The ideal model system for the evolution of Al tolerance in wild grasses is one which has strong tolerance but also harbors functional genetic variation, providing a basis by which to later dissect the trait at the genetic level. In the literature, intraspecific variation in Al tolerance has been documented only rarely in wild grasses (Taylor *et al.* 1966; Wheeler *et al.*, 1992b; Wheeler, 1995; Kidd & Proctor, 2001). However, by conducting field surveys in old-fields in upstate New York in 2009 I identified two grass species that occur regularly across a wide range of soils in the region from pH 6.7 to 4.5 with Morgan extractable Al content from 10 to 300 ppm (unpublished data): *Dactylus glomerata* (orchard grass) and *Anthoxanthum odoratum*. Of the two, *A. odoratum* occurs in much higher numbers (pers. obsv.). *A. odoratum* has been the subject of ecological research (Antonovics & Bradshaw, 1970; Berendse, 1983; Antonovics & Schmitt, 1986; Platenkamp & Shaw, 1992; Silvertown *et al.*, 2005; Freeland *et al.*, 2010, 2011) but is perhaps best known from early studies of local adaptation among populations growing on contaminated mine tailings in Europe (McNeilly & Antonovics, 1968; Antonovics, 1972). In these and at least one other pair of studies (Qureshi *et al.*, 1985; Al-Hiyaly & Bradshaw, 1990) it has been tested experimentally for tolerance to heavy metals and has been studied in the context of Al tolerance (a "light" metal) in one location, the long term agricultural Park Grass Experiment in Harpenden, UK (Davies & Snaydon, 1973). The species is a biennial to perennial pasture grass that is primarily outcrossing and was introduced to North America from Europe in the 1700's (Pitcher *et al.*; Grant *et al.*, 1978).

For herbaceous plants such as *A. odoratum,* metal tolerance can be assessed in one of two ways: using sand culture or hydroponics. Tolerance can also be quantified using a variety of metrics including relative root growth, root biomass, shoot biomass, and grain yield (Howeler, 1991). Because Al causes damage to plants primarily by damaging growing root tips (Kochian, 1995), measurement of relative root growth (RRG) in hydroponic culture has become the most common method used in studies of Al tolerance in cultivated plants. Hydroponics also typically offers more precise control of

Al exposure levels in comparison with sand culture methods (J. Shaff, pers. comm.). In this method, either young seedlings, or replicate clonal tillers are placed in a hydroponic growing solution containing either aqueous calcium nitrate or a more complete mixture of macro and micronutrients for growth. The length of the longest root is measured over several days, and then metal salts are added to the solution and root growth is measured again over an equal amount of time. The sensitivity of a plant genotype is measured as the absolute or fractional (percentage) difference in growth during the control and metal exposure time periods (= relative root growth, RRG). Replicate tillers from each plant can be tested to get an average RRG for each plant genotype. As an added control, RRG for each genotype can be measured in solution without aluminum for comparison.

Control of experimental conditions is critical in metal tolerance studies in order to draw meaningful conclusions, however these controls are commonly overlooked (J. Shaff, pers. comm., see also discussion of this topic in Kochian *et al.,* 2004, Famoso *et al.* 2010). First, it is important to take into account the activity level (i.e. biological availability) of metals and nutrients in hydroponic solution, which are often quite different than the base concentration due to chemical complexing (Shaff *et al.*, 2009). The activity level of a metal salt added as a stressor may be quite different than the physical concentration in solution, and the presence of a metal may significantly alter the availability of specific micro and macronutrients when added to solution. In particular, Al complexes at a high rate with phosphate, sulfate, and iron in solution, depending on concentration. Low solubility Al hydroxides can also form quickly depending on pH and precipitate out of solution. Second, in order for RRG to be a meaningful measure of metal tolerance care must be taken either to verify that plants are measured while in a relatively linear phase of root growth, or that RRG is measured for comparison in control solution. Third, the overall strength (toxicity) of the added metal must be calibrated so that exposure is just strong enough to detect variation in tolerance between genotypes, but does not leave the majority of genotypes either unaffected or completely halted in growth. Lastly, it is important to consider differences in tolerance between plant life stages. For example, young seedlings may be more or less sensitive to metals and nutrients and this difference may be a critical component of overall plant fitness.

With these considerations in mind, I modified existing hydroponic metal tolerance testing protocols for specific use in measuring both seedling and adult (tiller) Al tolerance in *A. odoratum*. Here I present protocols for collection, cultivation, and phenotypic characterization of Al tolerance, and use them to demonstrate that variation in soil Al tolerance exists within the species. These protocols provide tools necessary for future ecological genetic work on Al tolerance in the wild, and provide a means for making evolutionary genetic comparisons across wild and cultivated grasses.

Methods

Plant and seed material. In summer of 2009 I sampled *A. odoratum* plants from six old-fields that span a range of pH and Al content in the vicinity of Ithaca, NY (Table1). I identified plants by their inflorescences (Fig. 1) and removed them by shovel with intact field soil in June and early July 2009. I brought them back to the Cornell University greenhouse where I removed most of the surrounding plants and field soil and repotted each plant in standard greenhouse soil. In August I transplanted newly formed tillers to 10 inch pots filled with a 1:1 mixture of vermiculite and perlite. I grew them in the greenhouse at approximately 25 deg C days, 15 deg C nights, with supplemental metal-halide lighting. Plants were treated with 21-5-20 NPK fertilizer at 150 ppm twice per week for several weeks, and once established, with 21-7-7 NPK fertilizer once every 1-2 weeks. I trimmed back the leaves intermittently to promote healthy growth. To control occasional infection by powdery mildew plants were treated periodically with fungicide.

In July 2009 I also collected seeds from each of the six fields. I collected entire inflorescences every 10 paces along 2-3 parallel transects spaced 20 paces apart in each field. Seeds consist of an enclosed caryopsis attached to an awned "hull" (Fig. 1, see also Antonovics and Schmitt, 1986). Awns rotate hygroscopically. The largest number of ripe seeds came from inflorescences that contained a mixture of both brown and green florets. I removed seeds from the seed heads in the lab by placing them flat on paper and manually scraping them with a metal spatula. Upon separation some seeds remained attached to their hulls while others did not. Once separated from the flowers, seeds were

stored at room temperature in 1.5 mL tubes with a small hole poked in the caps for ventilation.

Rice growing solution assay. In the most Al tolerant cereal crop, rice, Al tolerance is measured in a high nutrient concentration hydroponic solution containing 160 uM activity Al^{3+} in solution (Famoso *et al.*, 2010) (Table 1). To test the efficacy of this same assay on *A. odoratum* I placed two tillers from each of 8 plants in this same solution. I trimmed one tiller of each pair of all attached roots and left the other intact, containing a few to many soil-established roots. I placed tillers into slots cut in 1/8 inch thick foam sheets, floated in two 16 L opaque plastic tubs each containing 8 L of rice growing solution with aeration from a single plastic tube attached to an aquarium pump (Aqua Culture). The tubs were kept in a growth chamber at 20 deg C day/ 18 deg C night, 13 hour days. Tillers were grown for 14 days and on day 15 the hydroponic mixture was replaced with one containing standard 160 uM activity Al3+. The length of the longest root (whether old or newly generated) was recorded throughout, ending on day 30. Tubs and other equipment were acid washed in 2% hydrochloric acid in between trials to remove any residual aluminum.

Modified Long-Ashton solution assays. Davies and Snaydon (1973) tested Al tolerance of *A. odoratum* using a mixture similar to a diluted Long-Ashton solution (Hewitt, 1966). I also tested *A. odoratum* in two modifications of the Long-Ashton (LA) mix (Table 2). LA solution 1 contained macronutrients at 1/3 original strength and solution 2 at 1/10 original strength. In both solutions I included micronutrients at full strength to avoid exhaustion during the growing period of the assay. I also included a very low amount of phosphorus because phosphorus is known to complex at a high rate with Al in solution (and in natural soils, Rengel, 2003). *A. odoratum* has been shown to be minimally sensitive to lack of P (Davies and Snaydon, 1973) so I designed the solution with very little P in order to minimize availability differences between the control and treatment solution containing Al. As it is unknown the form in which *A. odoratum* most efficiently utilizes nitrogen, I included nitrogen as 1/3 ammonium and 2/3 nitrate (Table 1). I mixed large batches of solution in clean 120 L plastic trashcans (Rubbermaid) and split the mix

evenly into tubs for testing. The solutions were slowly adjusted to a pH of 4.5 using 0.7% HCl. I used three 10 L capacity opaque plastic tubs filled with 9L of solution to test tillers: two with the 1/3x modified LA mix and one with the 1/10x modified LA mix. I aerated each tub as in the rice-solution trial but with the addition of a silica (rather than plastic or calcium composed) airstone diffuser at the end of each aeration tube (Aquatic Ecosystems, Inc.).

I separated three tillers from each of 18 greenhouse plants, trimmed off all roots and placed one in each tub. I placed tillers in slots cut into in floating rafts this time made of closed-cell polyethylene foam (cat no. 8726K23, McMaster Carr, Robbinsville, NJ). I marked the solution level with tape on each tub and added distilled water daily to maintain volume, countering losses due to evaporation and transpiration. Growth of tillers in the hydroponics tended to mildly acidify the solution, possibly due to nitrate uptake (McClure *et al.*, 1990), so I monitored pH of each tub daily and adjusted the solution to 4.5 with 1M NaOH when necessary. I measured the length of the longest root for each tiller over a period of 18 days, changing the solutions once with regular solution at 6 days and then with treatment solution at 12 days. One tub of 1/3x solution was used as a control and contained no Al in the treatment solution. Treatment solutions for the other two tubs contained nutrients plus 300 uM activity Al.

I quantified Al tolerance of each tiller as the difference between the amount of root growth from day 6 to 12 (no Al) vs. day 12 to 18 (Al added). (In the control treatments the solution is replaced but no Al is added at day 12.) Lower (more negative) relative root growth (RRG) indicates lower Al tolerance in the Al exposure treatments.

Control of nutrient and Al activities. I considered chemical complexing in solution carefully in the design of the hydroponic mixes. I used the program GeoChemEZ (Shaff et al., 2009) to predict the activity of all nutrients in solution with the goal of designing a solution that reduces nutrient availability differences between the control and Al treatment solutions. Simulations were run over 75 or more iterations with pH fixed at 4.0 or 4.5. Based on the results I added Al salt (as aluminum chloride heptahydrate) to the 1/3x and 1/10x modified Long Ashton solutions at a final concentration of 1,200 and 950 uM respectively, to achieve an Al^{3+} activity level 300 uM in both (Table 2).

Seed germination. I tested several methods of germination for *A. odoratum* seeds including planting on standard greenhouse soil, vermiculite, filter paper, rolled rice germination paper (Anchor Paper, St. Paul, MN), pellet filled germination cups suspended in hydroponic solution, and on gel nutrient media. I tested from 20 to 90 seeds in each treatment (Fig. 4) and recorded germination for approximately 16 days. For each test, I used seed from three or more of the local populations and placed test seeds in in a growth chamber on 12-13 hour days, 16 degree C night/ 20 degree C days. I tested the effect of cold exposure (vernalization) on germination in combination with some treatments by exposing to the seeds to 4-7 days of 10 degree, low light, 8 hour days, just after planting. I also looked for germination differences between seeds that were planted with hulls attached vs. not attached (Fig. 1).

To generate seedlings for hydroponic testing, seeds must be germinated in relatively sterile conditions and must be transplanted to the hydroponics with intact roots. I was able to accomplish this by germinating seeds on solidified hydroponic media. I created the media by adding ultrapure agarose at 1% w/v to 1/10x LA solution, sterilizing the mixture in an autoclave, and then aliquoting 8-10 mL into 15 mL conical tubes (BD Biosciences). I separated seeds into round-bottomed 96 well plates, sterilized the seeds in the plates with 10% bleach for 10 minutes and then rinsed three times with distilled water. I then soaked the seeds in distilled water overnight. I placed three to five seeds in each media-filled tube and covered them individually covered with Parafilm. I also tested some seeds on media that I acidified (to pH 4.5) prior to autoclaving it (intrinsic solution pH = 5.4), and tested the effect of gibberelic acid on germination by soaking seeds in 500 ppm GA for 7 hours prior to planting.

Seedling Al tolerance experiment. I tested seedlings between 5 and 14 days old (that had germinated either on the nutrient media or in germination paper rolls) in hydroponic culture similar to the adult tillers. I placed eight to ten seedlings with roots intact in rafts in each of two 10 L tubs of aerated 1/10x LA solution. After 10 days of growth, I replaced the solution in tub 1 (control) with fresh control solution and tub 2 (aluminum treatment) with solution containing 160 uM activity Al^{3+}. At day 16 I replaced the

solutions again, this time tub 2 with one containing 300 uM Al^{3+} activity, tub 1 remaining control.

Results

Plant Cultivation. In order to reduce root damage and water stress, I used a shovel to remove whole plants from the field with intact roots and a substantial amount of field soil. The procedure was arduous, but plants remained healthy throughout, showing little sign of stress. When fully separated from the field soil and transplanted to potting soil in the greenhouse, plant growth was vigorous and often produced tightly rolled leaves, which may be caused by transfer to a high nutrient environment (a syndrome called "buggy whipping" in corn). Growth was much more moderate, producing normally formed leaves, when tillers were transplanted to a 1:1 perlite/vermiculite soil-less mixture, with nutrients supplied as dilute fertilizer periodically (see methods).

Germination of field collected seed also proved a successful way to generate *A. odoratum* plants at the seedling stage. Germination rates were highest in moist vermiculite (50%) followed by soil (up to 37%) (Fig. 2). Germination was slightly improved with a brief chilling period after planting. I also found that seeds attached to the hull germinated at only a slightly lower rate than un-hulled seeds. Seeds also germinated well, up to 35%, on solidified hydroponic solution, and seedlings were easily removed with roots intact using tweezers, and placed in the hydroponics. Seeds germinated at a slightly higher rate when the solid media was acidified, and when they were exposed to gibberelic acid prior to planting.

Hydroponic assay in rice growing solution. Tillers from the greenhouse regenerated easily when separated from the parent plant and grown in the rice hydroponic solution. In solution (without added Al), tillers generated new, rapidly growing roots within the first few days after transplant, while old soil-established roots only continued to grow in solution for about half the tillers tested (Fig. 3). After 15 days growth most tillers had extended their initial root length by 5-10 cm. Upon exposure to the 160 uM activity Al solution, root growth in some tillers began to slow but for most continued strongly until

after about 6 days exposure. At least one tiller continued to grow without slowing until the end of the Al exposure period (15 days).

Hydroponic assays in Modified Long-Ashton solutions. I used GeoChemEZ to design two different modifications of the Long-Ashton solution for tolerance testing. The levels of nutrient availability between the control and Al containing solutions for these two mixes are not significantly altered by the addition of Al at an activity level of 300 uM (Fig. 4), except for phosphorus. Although 94% of phosphate in solution becomes complexed with Al in the treatment solution, the pre- and post- Al addition levels of available P were both near zero. In practice, when mixing the solutions, I observed some precipitation of Al out of solution in the 1/3x solution but not in the 1/10x solution.

When grown in both modified Long-Ashton solutions, tillers initiated new roots that reached lengths comparable to those in the rice growing solution over a 12 day period, and at the end of the experiment total root growth attained by plants in all three mixes was similar (Fig. 3 and 5). In the 1/3x control treatment, tillers grew an average of 1.1 cm less in period 2 (day 12-18) than in period 1 (day 6-12) reflecting a natural slowing of growth (Fig. 5A). However, in the Al treatments the slowing of root growth was doubled to a difference of 2.0 and 1.9 cm in the 1/3x and 1/10x solutions, respectively (Fig. 5B and 3C). Overall, root growth slowed more quickly under Al exposure at 300 uM in these treatments than it did in the rice solution containing 160 uM Al (Fig. 6).

General methodological considerations. Throughout these experiments with tillers (and some others for which data is not shown) I noted and addressed some important general methodological issues. First, measurement of the longest root was an easier and more accurate means of judging overall root growth than measurement of a single root on each plant that had been marked with pen or string. Second, unlike for rice aeration of the hydroponic solution was essential for growth and was best when airstones were used for diffusion. Third, in some trials a significant amount of brown film developed on some roots, which was most likely infection by a water mold, such as *Pythium* (Paul Cooper, pers. comm.). I was able to reduce infection by sterilizing all equipment with a dilute

solution of greenhouse cleaner GreenShield (BASF Inc.) in between trials and by sterilization of tiller bases for 30 seconds in 10% bleach followed by a distilled water rinse before placing them in the rafts. Some growth of algae was also noted, particularly with use of the high nutrient rice solution, but could be controlled by replacing the nutrient solution with freshly made solution at a minimum every 7 days.

Seedling tolerance testing. I tested the growth and Al tolerance of seedlings in the 1/10 strength LA solution. Seedlings, particularly over 3 cm in shoot height when transplanted, grew well in the hydroponic assay with only moderate signs of transplant stress that generally disappeared after several days acclimation. Without Al exposure, seedling root growth was fairly linear over a period of 10 days with a slight increase in growth by 16 days (light grey lines, Fig. 7A**,** control RRG measured in 7B and 7C). Exposure to 160 uM Al activity did not appreciably effect root growth rate over six days (Fig. 7B). When exposure was increased to 300 uM Al, seedling root growth rate slowed appreciably (Fig. 7C).

Discussion

Overall, *Anthoxanthum odoratum* has proved a tractable species for study in the greenhouse, growth chamber, and in hydroponic root growth assays. It has high resilience to transplant from the field, grows well in soilless media, and has a high capacity for rapid root regeneration. Seeds are abundant and germinate in vermiculite, soil, or sterile growth media at moderate rates (often 30% or more). Seeds germinate readily without time-consuming separation from their enclosing hulls, and once germinated seedlings can be transferred from a solid growing medium into hydroponic growth conditions with minimal stress. Genotypes can also be phenotyped for Al tolerance with high precision through measurement of replicate clonal tillers or large seed families in hydroponic culture. Tillers are extremely resilient and quickly regenerate roots that can be assayed under Al exposure conditions.

The 1/10x LA solution used in these experiments proved the most effective way to accurately test Al tolerance in this species, more so than a solution typically used to

test the tolerance of rice, and more so than a similar solution with higher nutrient concentrations. The 1/10 solution more closely simulates the level of nutrients found in acidic field soils than most other solutions used in the literature, and allows for precise control of Al activity without precipitation or significant alteration of micro- and macronutrient levels between Al and non-Al treatments. Usefully, the same nutrient solution can be used to test the tolerance of *A. odoratum* in both seedlings and adults, and to create solid germination medium.

In these preliminary trials *A. odoratum* has proved surprisingly tolerant to Al, more so than most other herbaceous plants and other temperate grasses that have been tested to date (Wheeler *et al.*, 1992a; Poozesh *et al.*, 2007). The level of exposure used to test seedlings of the most tolerant crop grass, rice, was not sufficient to significantly affect root growth in *A. odoratum.* An activity of 300 uM Al, close to the limit of solubility in this and most other nutrient solutions, was required to slow root growth just enough to allow the measurement of tolerance across a small set of genotypes. Unexpectedly, young seedlings of *A. odoratum* proved just as tolerant if not more tolerant than adult tillers. Findings here also suggest that considerable genotypic variation exists for Al tolerance within the species.

Ecological and agricultural plant geneticists both stand to gain from integration of their research on Al tolerance in grasses, from discovery of novel mechanisms of soil Al tolerance in the wild that may point toward homologous targets for genetic engineering or breeding in cereals. *Anthoxanthum odoratum*'s experimental tractability, unique high level of Al tolerance among grasses, and considerable intraspecific functional genetic variation for the trait make it a promising system for the study of the genetic basis of this important plant trait in natural settings.

Acknowledgements. Special thanks to Paul Cooper for help and advice with plant cultivation; Jon Shaff, Eric Craft, Adam Famoso, and Leon Kochian for help and advice in hydroponic development; Peter Marks and Tim Levitich for identification and access to local plant populations; members of the Geber and McCouch labs for support

throughout. This work was supported by funding from the Cornell Chapter of Sigma Xi, the Department of EEB, and the BEB Small Grants Program.

FIGURES

Figure 1. *Anthoxanthum odoratum* inflorescences (right) at two stages; early flowering (lower) and ripening (upper). *Anthoxanthum* seeds separated from the hull (scale bar is 2 mm).

Figure 2. Seed Germination Treatments

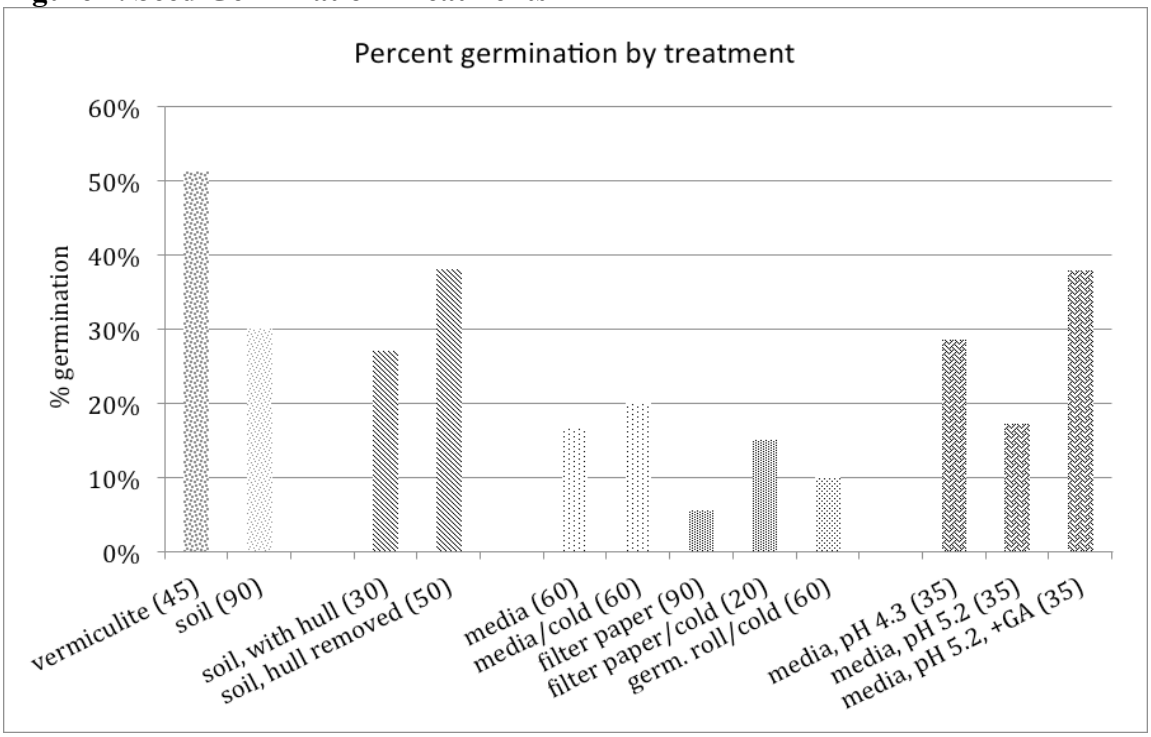


Figure 3. Growth trajectories of 12 tillers growing in rice growing solution. Al was added to an activity of 160 uM on day 15 (dotted line). Grey lines are tillers with roots trimmed to starting length of 0.5 cm; black lines roots left un-trimmed.

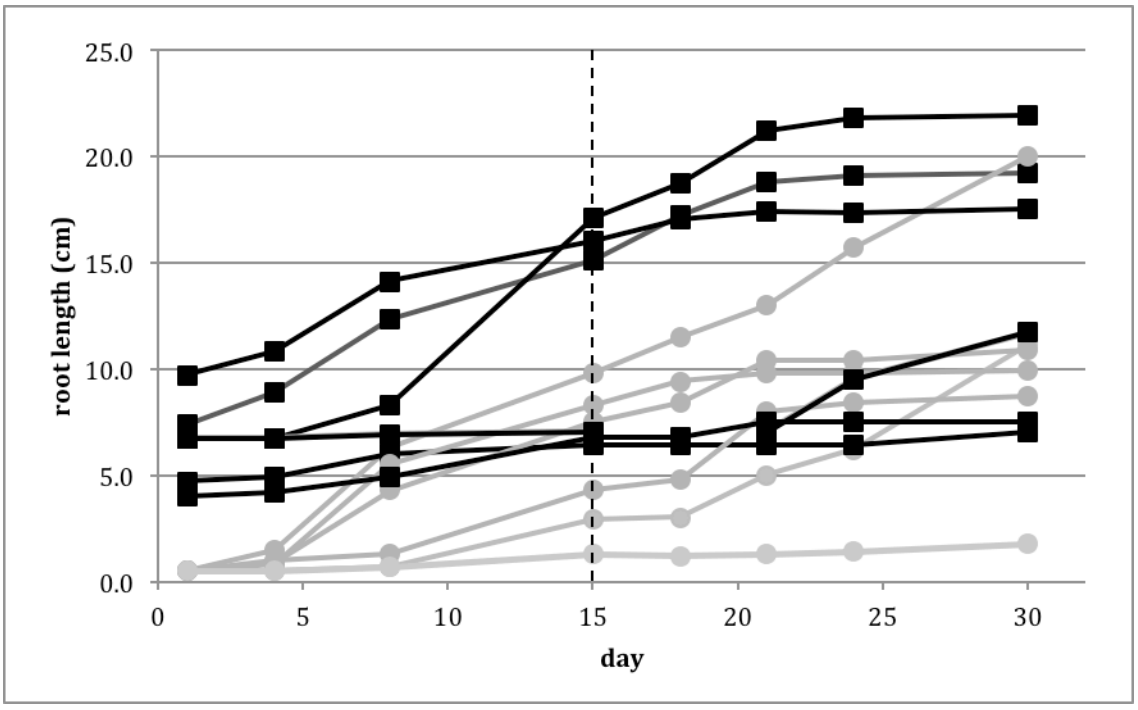
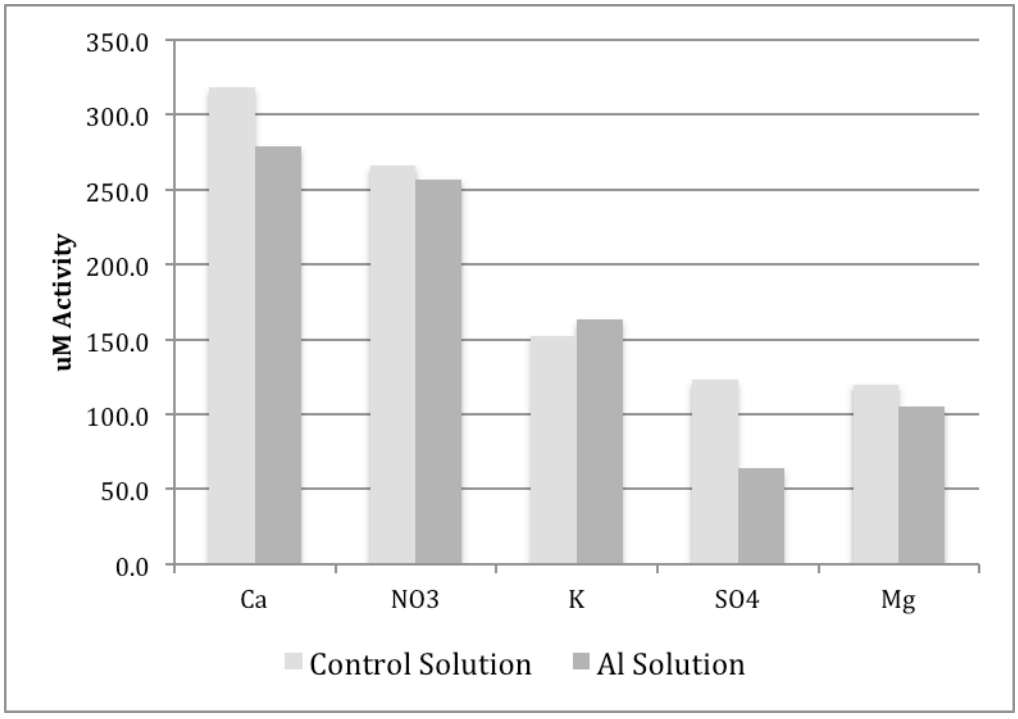


Figure 4. Nutrient activity levels in 1/10x modified Long-Ashton solution. Macronutrients (top) and micronutrients (bottom). Light grey bars, without Al; dark bars, with Al at 950 uM concentration.



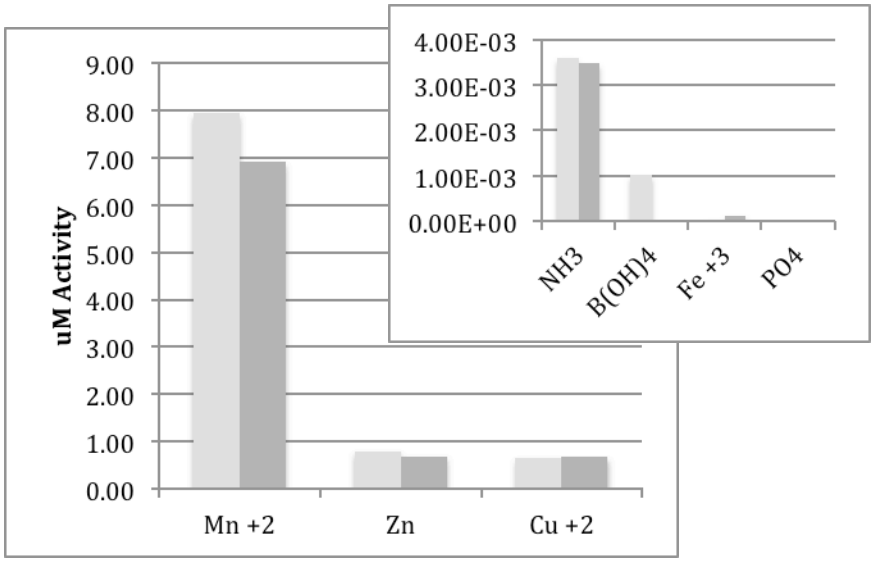
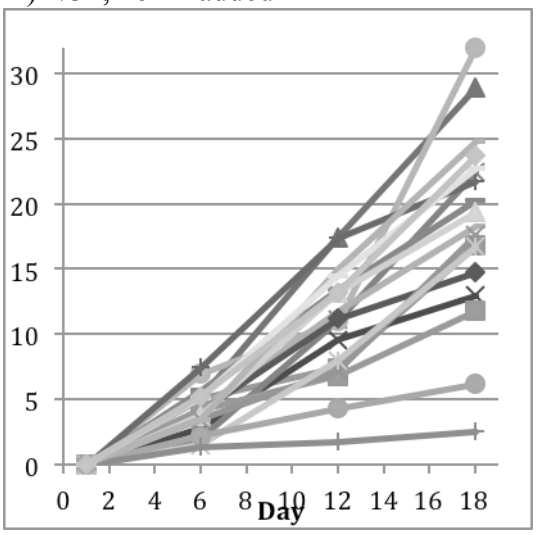
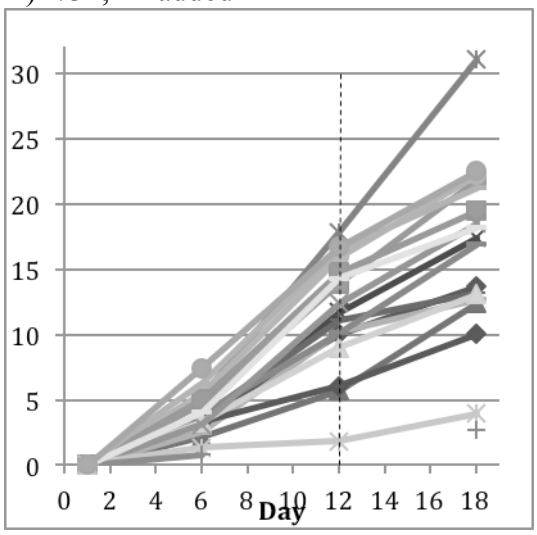


Figure 5. Growth trajectories of individual tillers growing in control and aluminum exposure treatments in 3 growth solutions. Each line represents one genotype. (A) 1/3 strength modified LA soltuion, no Al exposure (control); (B) 1/3 strength modified LA soltuion, Al exposure (dotted line); (C) 1/10 modified LA solution with Al exposure. Dotted line represents the day on which Al treatment was implemented (see methods).

A) 1/3x, no Al added



B) 1/3x, Al added



C) 1/10x, Al added

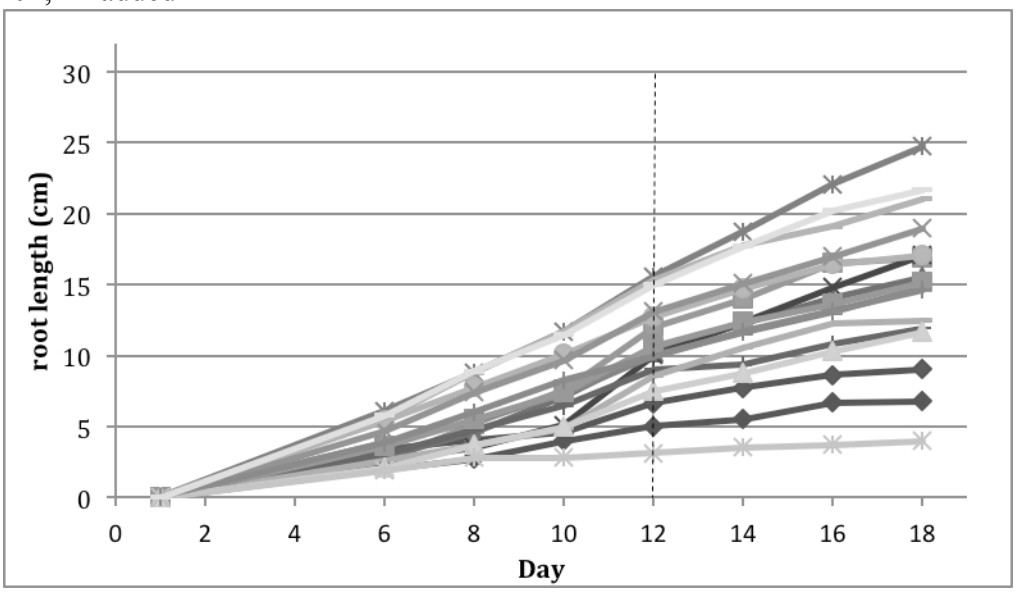


Figure 6. Six day relative root growth (RRG) of tillers growing in three hydroponic treatments: 1/3 strength, modified LA soltuion without Al (control); 1/3 strength modified LA soltuion, Al exposure; and 1/10 strength, modified LA solution with Al exposure, respectively (see methods). Mean RRG of box plots separated by different letters are significantly different than each other at $p<0.05$ (Tukey's HSD). Al was added to 300 uM activity.

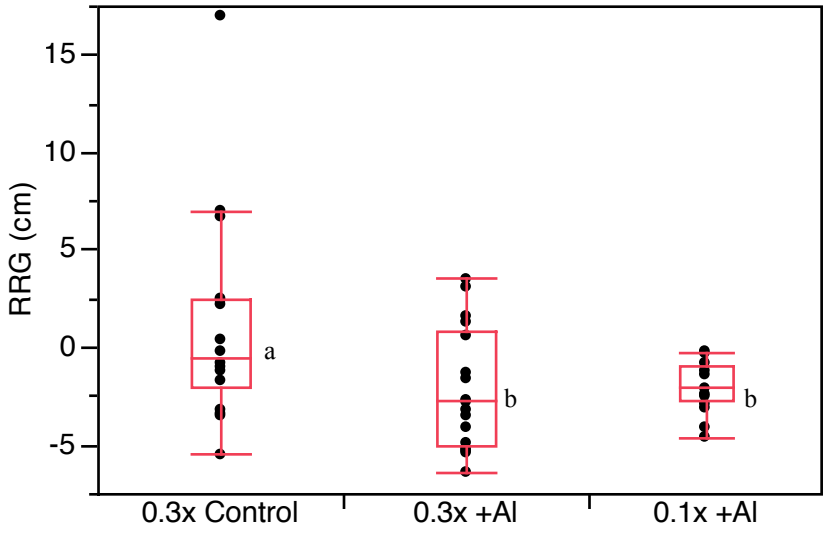
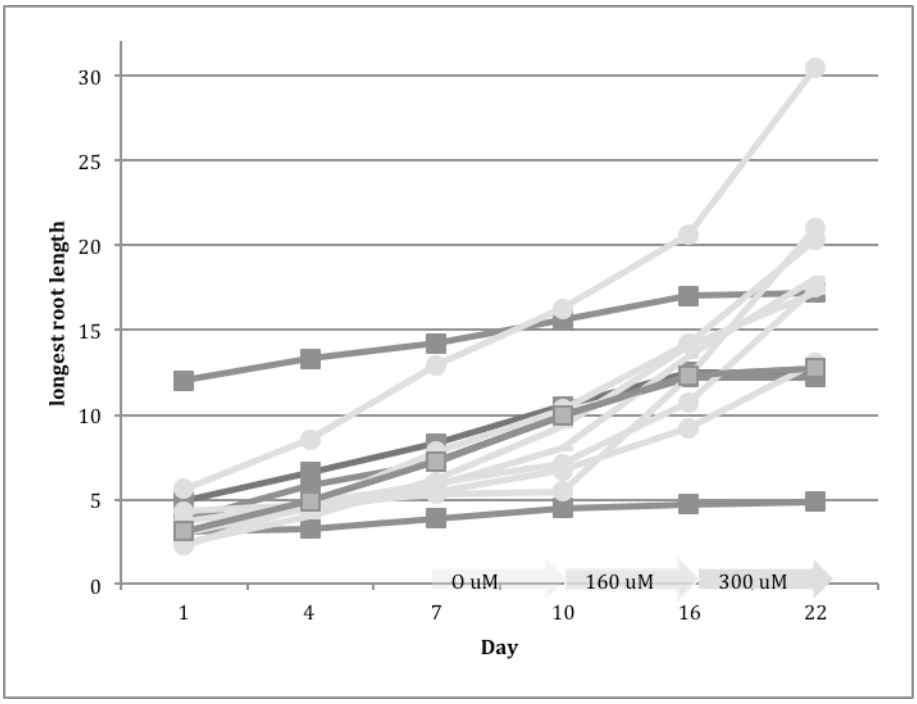
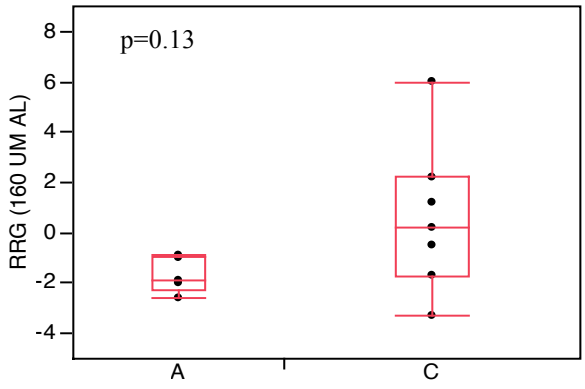


Figure 7. A) Growth trajectories of 10 seedlings. Dark lines are seedlings in the Al treatment. Lighter lines are in the control treatment (no Al). Arrows along the x-axis represent replacement of the growing solution with nutrient solutions containing the indicated activity of Al^{3+} in the Al treatment group. Three seedlings died in each treatment (not shown). B) Comparison of RRG between the 0 uM Al and 160 uM Al treatment periods. C) Comparison of RRG across control (C) and Al (A) treatments.

A)



B)



C)

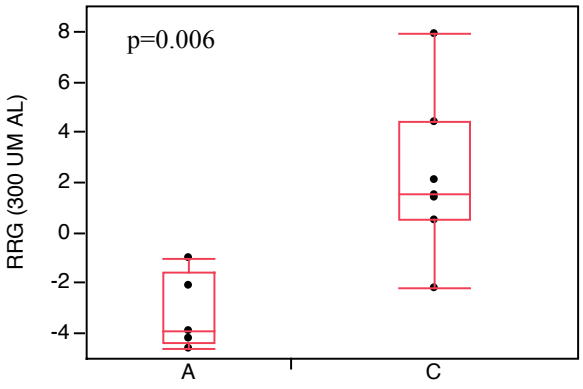


Table 1. Relative soil element levels from six field sites where *A. odoratum* was collected. Values are in mM concentration, estimated from Modified Morgan (sodium acetate) extracts of field soils*.

field site	pH	Al	P	K	Mg	Ca	Fe	Mn	Zn
Bailor	4.5	10.6	62.4	916.5	548.5	3874.7	719.2	813.0	17.8
Hanshaw3	4.9	7.8	25.0	1321.4	1268.3	7211.3	1889.0	788.7	12.2
Long Point	5.1	9.3	25.0	1385.3	2433.8	14153.5	456.6	388.3	14.0
Hanshaw 2	5.3	6.9	12.5	895.1	2056.8	12861.9	552.1	130.4	14.8
Hanshaw 1	6.1	0.8	12.5	831.2	9941.0	73619.6	65.7	81.9	6.6
Turkey Hill	6.4	0.6	33.3	959.1	12066.4	82660.6	44.8	45.5	4.6

* see Bertsch and Bloom, 1996.

Table 2. Concentration of nutrients and aluminum in hydroponic solution.

	Rice	Long-Ashton[1] (LA)	Davies-Snaydon[2]	1/3x modified LA	1/10x modified LA
	mM	mM	mM	mM	mM
Ca	1	4	0.24	1.2	0.4
K	1.05	4	0.08	1.2	0.4
Mg	0.86	1.5	0.02	0.45	0.15
NO3	2.5	12	0.32	0.84	0.28
NH4	1.5	0	0	0.36	0.12
SO4	0.2	1.5	0.29	0.45	0.15
PO4	0.05	1.3	0.03	0.05	0.05
	uM	uM	uM	uM	uM
Fe	77.0	50	1.4	77.0	77.0
B	33.0	50	0.6	50.0	50.0
Mn	11.8	10	0.33	10.0	10.0
Cu	0.8	1	0.026	1.0	1.0
Mo	1.1	0.5	0.011	0.5	0.5
Zn	3.1	1	0.023	1.0	1.0
Al conc.	540	NA	2000	1200	950
Al activity	160	NA	429	~300	300

Literature Cited

Al-Hiyaly SAK, Bradshaw AD. **1990**. The effect of zinc contamination from electricity pylons. Contrasting patterns of evolution in five grass species. *New Phytologist* **114**: 183–190.

Antonovics J. **1972**. Population Dynamics of the Grass Anthoxanthum Odoratum on a Zinc Mine. *Journal of Ecology* **60**: 351–365.

Antonovics J, Bradshaw AD. **1970**. Evolution in closely adjacent plant populations VIII. Clinal patterns at a mine boundary. *Heredity* **25**: 349–362.

Antonovics J, Schmitt J. **1986**. Paternal and maternal effects on propagule size in *Anthoxanthum odoratum*. *Oecologia* **69**: 277–282.

Berendse F. **1983**. Interspecific competition and niche differentiation between *Plantago lanceolata* and *Anthoxanthum odoratum* in a natural hayfield. *Journal of Ecology* **71**: 379–390.

Bertsch, PM., Bloom, PR. 1996. Methods of Soil Analysis: Part 3. Chemical Methods. Madison: Soil Science Society of America and American Society of Agronomy.

Davies MS, Snaydon RW. **1973**. Physiological differences among populations of *Anthoxanthum odoratum* L. collected from the Park Grass Experiment, Rothamsted. II. Response to aluminium. *Journal of Applied Ecology* **10**: 47–55.

Famoso A, Clark R, Shaff J, Craft E, McCouch SR, Kochian L. **2010**. Development of a novel aluminum tolerance phenotyping platform used for comparisons of cereal Al tolerance and investigations into rice Al tolerance mechanisms. *Plant Physiology.* **153**: 1678-1691.

Foy CD. **1988**. Plant adaptation to acid, aluminum-toxic soils. *Communications in Soil Science and Plant Analysis* **19**: 959–987.

Freeland JR, Biss P, Conrad KF, Silvertown J. **2010**. Selection pressures have caused genome-wide population differentiation of *Anthoxanthum odoratum* despite the potential for high gene flow. *Journal of Evolutionary Biology* **23**: 776–782.

Freeland JR, Biss P, Silvertown J. **2011**. Contrasting patterns of pollen and seed flow influence the spatial genetic structure of sweet vernal grass (*Anthoxanthum odoratum*) populations. *Journal of Heredity* **103**: 28–35.

Grant MC., Janis Antonovics, Antonovics J. **1978**. Biology of ecologically marginal populations of *Anthoxanthum odoratum*. I . Phenetics and dynamics. *Evolution* **32**: 822–838.

Haling RE, Simpson RJ, Culvenor RA, Lambers H, Richardson AE. **2010**. Effect of soil acidity, soil strength and macropores on root growth and morphology of perennial grass species differing in acid-soil resistance. **34**: 444-456.

Hewitt EJ. **1966**. Sand and water culture methods used in the study of plant nutrition. *Tech. Commun. Commonw. Bur. Hort. Plant Crops,* No. 22 (2nd edn.)

Howeler RH. **1991**. Identifying plants adaptable to low pH conditions. Plant Soil interactions at low pH. Amsterdam, The Netherlands: Kluwer, 885–904.

Kidd PS, Proctor J. **2001**. Why plants grow poorly on very acid soils: are ecologists missing the obvious? *Journal of Experimental Botany* **52**: 791–9.

Kochian L V. **1995**. Cellular mechanisms of aluminum toxicity and resistance in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **46**: 237–260.

Kochian, Leon V, Hoekenga, O, Pineros, M. 2004. How do crop plants tolerate acid soils? Mechanisms of aluminum tolerance and phosphorous efficiency. *Annual Review of Plant Biology*. **55**: 459-93

McClure PR, Kochian L V, Spanswick RM, Shaff JE. **1990**. Evidence for cotransport of nitrate and protons in maize roots. II. Measurement of $N0^{3-}$ and H^{+} fluxes with ion-selective microelectrodes. *Plant Physiology*: 290–294.

McNeilly T, Antonovics J. **1968**. Evolution in closely adjacent plant populations IV. Barriers to gene flow. *Heredity* **23**: 205–218.

Pitcher D, Russo MJ, Conservancy TN. **1988**. Element stewardship abstract for *Anthoxanthum odoratum*, sweet vernal grass. Arlington: The Nature Conservancy.

Platenkamp GAJ, Shaw RG. **1992**. Environmental and genetic constraings on adaptive population differentiation in *Anthoxanthum odoratum*. *Evolution* **46**: 341–352.

Poozesh V, Cruz P, Choler P, Bertoni G. **2007**. Relationship between the Al resistance of grasses and their adaptation to an infertile habitat. *Annals of Botany* **99**: 947–54.

Qureshi JA, Thurman DA, Hardwick K, Collin HA. **1985**. Uptake and accumulation of zinc, lead and copper in zinc and lead tolerant *Anthoxanthum odoratum* L. *New Phytologist* **100**: 429–434.

Rao IM, Zeigler RS, Vera R, Sarkarung S. **1993**. Selection and breeding for acid-soil tolerance in crops: upland rice and tropical forages as case studies. *BioScience* **43**: 454–465.

Rengel Z. **2003**. *Handbook of Soil Acidity*. New York: Marcel Dekker Inc.

Schonfeld MA, Chancellor RJ. **1983**. Factors influencing seed movement and dormancy in grass seeds. *Grass & Forage Science* **38**: 243–250.

Shaff J, Schultz B, Craft E, Clark R, Kochian L. **2009**. GEOCHEM-EZ: a chemical speciation program with greater power and flexibility. *Plant and Soil* **330**: 207–214.

Silvertown J, Servaes C, Biss P, Macleod D. **2005**. Reinforcement of reproductive isolation between adjacent populations in the Park Grass Experiment. *Heredity* **95**: 198–205.

Taylor P, Liu H, Heckman JR, Murphy, JA. 1996. Screening fine fescues for aluminum tolerance, *Journal of Plant Nutrition,* **5**: 677-688.

Wenzl P, Patiño GM, Chaves a L, Mayer JE, Rao IM. **2001**. The high level of aluminum resistance in signalgrass is not associated with known mechanisms of external aluminum detoxification in root apices. *Plant Physiology* **125**: 1473–84.

Wheeler DM. **1995**. Relative aluminium tolerance of ten species of Graminae. *Journal of Plant Nutrition*. **11**: 2305–2312.

Wheeler DM, Christie RA, Gardner R. **1992a**. Effect of aluminium on the growth of 34 plant species: A summary of results obtained in low ionic strength solution culture. *Plant and Soil* **146**: 61–66.

Wheeler DM, Edmeades DC, Christie R a. 1992b. Effect of aluminum on plant chemical concentrations in some temperate grasses grown in solution culture at low ionic strength. *Journal of Plant Nutrition* **15**: 387–402.

CHAPTER 3

Regulation of canonical and novel genetic pathways is linked to soil aluminum tolerance in the highly tolerant wild grass, *Anthoxanthum odoratum*

Abstract

One fundamental question in evolutionary biology is whether phenotypic evolution occurs more often by the same or different genetic mechanisms. Few traits have been examined in enough detail at the genetic level in a diverse array of species in order to address this question. This is true especially of stress tolerance traits in plants which, like most ecologically important traits, are complex and multigenic in nature. Here we have begun to examine the genetic basis of soil Al tolerance in a wild grass, *Anthoxanthum odoratum* using high-throughput RNA sequencing across experimental treatments and genotypes. We then compare our findings with genetic information on Al tolerance in several cultivated grasses to determine whether similar or different pathways and mechanisms likely contribute to this ecologically and agriculturally important trait. We find that some of the same families of genes are involved in Al tolerance in this exceptionally aluminum tolerant wild species and the cultivated crop grasses. In addition we identify a suite of loci with no known sequence homologies among grasses or plants that are clearly involved in the Al stress response and may comprise an additional novel mechanism of tolerance.

Introduction

CHAPTER 3

Exposure to high levels of soil aluminum (Al) is a common environmental challenge for plants. Al is the third most abundant element in the earth's crust, present in most soils, but in acidic soils it forms mononuclear cations (Al^{3+}) that are phytotoxic even at low concentration (Rengel, 2003). Acidic, high Al soils are common throughout the world due to soil age and natural weathering, but also due to air and water pollution or intensive use of nitrogen based fertilizers in some areas. It is estimated that acidic soils (pH < 5.5) cover 40% or more of the arable lands on earth. Al poses major limitations to crop production, particularly in tropical Asia and South America, and Al tolerance has been studied intensively in crops (Kochian *et al.*, 2005). However, it is also an ecologically important trait in wild plants that form natural communities on acid soils. Most cultivated plants have low to moderate Al tolerance (Famoso *et al.*, 2010), while some, perhaps many wild species far surpass crops in Al tolerance. The diversity and evolution of Al tolerance mechanisms among diverse wild species has been little studied at the genetic level. Because there is growing information on the genes that confer Al tolerance in cereal crops, there now exists great opportunity for comparative exploration of the genetic basis of Al tolerance in related but non-cultivated grasses.

Much is known about the biology and genetics of soil Al tolerance in grasses that have a long history of selective breeding. Loci underlying QTL for Al tolerance have been identified in sorghum (Magalhaes *et al.*, 2007), wheat (Raman *et al.*, 2005), barley (Furukawa *et al.*, 2007), rye (Magalhaes *et al.*, 2004), maize (Maron *et al.*, 2010), and rice (Famoso *et al.*, 2011; Chen *et al.*, 2012). The earliest effect of Al exposure in these plants is damage to the growing root tips. Al binds with components of root cell walls causing the roots to become rigid and suffer lesions as they expand through the soil (Kochian *et al.*, 2005; Horst *et al.*, 2010). Variation in resistance to this type of damage in some cereals has been linked to variation in either induced or

constitutive organic acid release at the root tip (Ma, 2000). Organic acids chelate Al^{3+} ions in the rhizosphere when excreted in the zone just outside the root cell wall, and in some cases variation in only one or a few organic acid transport genes determines a large portion of within-species variation in Al tolerance (Raman *et al*. 2005; Magalhaes *et al.* 2007). In other cases organic acid exudation does not account for all the variation in Al tolerance between cultivars (Chen *et al.,* 2012; Huang *et al.*, 2009; Piñeros *et al.*, 2005). In these cases, variation may be due to the action of other less well-characterized Al exclusion mechanisms such as direct alkalization of the rhizosphere or various forms of cell wall modification (Kochian *et al.*, 2005; Ma, 2007). In addition, some plants are known to take up Al and sequester it in inert forms in their tissues (Hiradate *et al.*, 1997; Shen & Ma, 2001), but this does not appear to be prevalent within the grass family (with the possible exception of low level sequestration in rice (Famoso *et al.*, 2011; Xia *et al.*, 2010)).

The question of how often phenotypic evolution occurs by the same or different genetic pathways within and between species is a long-standing question in evolutionary biology (Schluter *et al.*, 2004; Arendt & Reznick, 2008). With regard to soil aluminum, the genetic basis of tolerance appears to have many parallels across cultivated grasses and even across some non-grass species (Ma, 2000, 2007; Kochian *et al.*, 2005; Magalhaes, 2006). Across the small set of cultivated species that have been examined, it appears that different organic acid transport genes within two major gene families have independently been recruited to confer tolerance during the domestication process (Ryan & Delhaize, 2010). But the question remains whether the trait has evolved by different or parallel pathways in species that have much higher tolerance than these crops and are more likely to have a long evolutionary history of selection for Al tolerance in natural settings. Within the wild grasses, phenotypic testing has shown that *Dactylus glomerata*

Phalaris aquatica, *Microlena stipoides* (Foy, 1988; Haling *et al.*, 2010), and *Brachiaria decumbens* (Wenzl *et al.*, 2001) have high Al tolerance that is likely to surpass tolerance found in cultivated grasses. Early work also showed that some wild grasses exude organic acids into the rhizosphere (reviewed Jones, 1998), however we are unaware of any studies of the genetic basis of the trait in species not subject to artificial selection. Al tolerance may have evolved in wild species by selection on organic acid transport genes and their expression patterns, or quite possibly by entirely novel pathways and mechanisms of tolerance. In this study we used next-generation transcriptomics and comparative genetic information from cultivated grasses to explore the genetic basis of Al tolerance for the first time in one highly Al tolerant wild grass species, *Anthoxanthum odoratum* (sweet vernal grass).

Anthoxanthum odoratum is a temperate grass native to Europe and introduced in the parts of the United States (Hedberg, 1990). It is a tetraploid, locally common, and occurs sometimes at high density in open meadows, successional old-fields, and in disturbed areas (Pitcher *et al.*, 1988). It is best known from classic early studies of the evolution of zinc and lead tolerance in populations growing on contaminated mine tailings in Europe (McNeilly & Antonovics, 1968; Antonovics, 1972). *A. odoratum* has also been the focus of evolutionary studies of population differentiation in response to soil modification at the historic Park Grass Experiment (PGE, Harpenden, UK) (Snaydon, 1970; Snaydon & Davies, 1972, 1976; Davies & Snaydon, 1973, 1976). At PGE, adjacent soil plots have been treated with lime and nitrogen fertilizers for over 110 years and the acidity of those plots has diverged over time (Silvertown *et al.*, 2006). Studies of Al tolerance in *A. odoratum* have shown that populations growing on different soil plots are locally adapted to the level of Al in the soil (Davies & Snaydon, 1973, Gould *et al*, in prep.).

Natural genetic standing variation for Al tolerance in *A. odoratum* is partitioned between the plots providing a convenient system by which to explore the basis of the trait.

We analyzed the aluminum content of roots and shoots of diverse genotypes from PGE to examine the relationship between Al accumulation in tissues and tolerance. We then conducted deep RNA sequencing (RNA-Seq) of root tips from one Al sensitive and one Al tolerant plant genotype in a factorial Al treatment experiment to identify transcripts involved in the Al stress response and for comparison with Al tolerance loci from cultivated grasses. Lastly, we examined the types of transcripts with different magnitudes of expression change in response to Al exposure in tolerant and sensitive plants. These candidates provide a means for further study of the evolution of locally adaptive differences in *Anthoxanthum*.

Methods

Plant material. Inflorescences (seed families) were collected in July 2010 from the Park Grass Experiment (Rothamsted Research, Harpenden, UK). Seeds were collected from across 10 pairs of plots of varying acidity (pH 7.2 to 3.7) and aluminum content (0-836 uM extractable in 0.1 $CaCl_2$, data not shown). Each pair of plots has been treated yearly with one level of nitrogen fertilizer at PGE, with the first plot in the pair also having received lime to adjust soil pH to about 7, the second plot in the pair left un-limed allowing the soil to acidify to about pH 4 (due to the repeated fertilizer applications). Collection and propagation procedures were followed as outlined in Chapter 1. Seeds were planted in vermiculite, vernalized for 3 days and then germinated in a growth chamber. At 5 weeks, one seedling per seed family was transplanted to the greenhouse. Plants were watered periodically with dilute acid-special fertilizer to maintain

growth. They grew in the greenhouse for a period of 10 months before tillers were removed for Al tolerance testing and genetic analysis.

Al and nutrient contents of roots and shoots. We measured the Al content of roots and shoots in *A. odoratum* seedlings. Sixty-four seedlings were allowed to grow and acclimate in hydroponic solution over the course of 21 days, the solution replaced every 7 days, to generate a large number of roots per seedling. Al tolerance of the seedlings was measured as relative root growth upon exposure to 950 uM Al (300 uM activity) for four days. At the end of the treatment period, approximately 1 cm long root tips were collected and frozen in liquid nitrogen. To generate enough tissue volume for Al content testing, tips were bulked into 11 pools according to plant Al tolerance. Each pool contained root tips from 4-6 seedlings, the first sample containing tips from the six most Al sensitive seedlings, the second sample from the next six, etc. The remaining shoot and root tissue from the seedlings was pressed and dried in an oven at 40 deg C for approximately 48 hours. 29-33 mg of dried tissue from the base of a mature leaf was cut from each dried plant and these were bulked into 11 pools in the same manner as the root tips. The Al and nutrient content of the shoot samples was analyzed by inductively-coupled plasma spectrophotometry (ICP) following acid digestion. The root tissue was thawed and separated into two fractions for ICP analysis, the cell wall fraction and the symplast (internal cell contents, mostly vacuolar sap), following the method of (Xia *et al*. 2010).

RNA sequencing. We measured Al tolerance of 108 plant genotypes using relative root growth as a measure of tolerance (Chapter 1) and then selected one tolerant and one sensitive individual from within one pair of plots at PGE for RNA-sequencing (Fig. 1). We chose plants from soils

treated with the same fertilizer but different pH as they are more likely to have genetic diversity at Al tolerance-related loci while retaining some homogeneity of genetic background. Nineteen tillers from the sensitive and the tolerant genotype each were separated from greenhouse plants and placed in each of 2 aerated hydroponic tubs. Tillers were grown without Al for 9 days and then on day 10 the solution in one tub was replaced with Al-free solution (1/10 strength Long-Ashton solution), the other with solution containing 950 uM $AlCl_3$. After 24 hours, approximately 1 cm long root tips were collected from all tillers and frozen in liquid nitrogen. Tissue was bulked into 4 groups (each made into a separate genetic library), one non-Al and one Al-treated sample for each of the two genotypes (Table 1). Total RNA was extracted from each bulk sample using Trizol (Invitrogen, Carlsbad, CA) followed by DNAse digestion. Preparation of RNA-seq libraries and sequencing was performed at the Weill Cornell Medical College using Illumina TruSeq paired-end sample preparation reagents and protocols. In this procedure mRNA is isolated from total RNA by poly-A selection, transcripts are chemically fragmented and then reverse transcribed using random priming. Complimentary DNA (cDNA) fragments were size selected on a gel resulting in sequences from 130 to 350 bp in length. Each library was individually barcoded and sequenced in parallel on the Illumina HiSeq2000 platform using 100 bp paired-end reads.

Transcriptome assembly. Initial output reads were de-multiplexed and filtered based on quality scores. Flow cell sectors with average base quality scores below 24 for any base position were discarded. The first 24 million reads from each of the four libraries was used to construct a composite *de novo* reference transcriptome using the assembler Trinity (Grabherr *et al.*, 2011) with a minimum contig length of 350 bp. The assembly was corrected for contig splitting and

lumping errors using the error correction module of the iAssembler (Zheng *et al.*, 2011) with parameters that combine contigs with greater than 98% similarity, at least 40 bp overlap, allowing 20 bp overhangs. To measure unigene expression, all filtered reads were then re-aligned to the transcriptome using TopHat (Trapnell *et al.*, 2012).

Functional annotation. A total of 88,103 unigenes were generated in the initial *de novo* transcript set and each was subjected to protein homology comparisons and gene ontology (GO) term mapping using the suite Blast2Go (Conesa *et al.*, 2005). Unigene sequence translations were compared with the GenBank nr protein database using blastx with an E-value cutoff of 10^{-6}. A small percentage of sequences had highest scoring matches to microorganisms (*Phytophthera*, 1,314; *Albug*o 236**,** others 65) and were eliminated from further analysis leaving 86,488 unigenes in the final transcriptome. We also conducted homology searches for specific Al tolerance loci from other species directly within the *A. odoratum* transcriptome using stand alone BLAST (tblastn). Unigene descriptions were generated based on significant BLAST matches using the BLAST Description Annotator in Blast2GO. We then assigned functional GO terms to a subset of described unigenes using confidence level cutoffs based on the default Blast2GO annotation configuration and evidence code weights (Conesa *et al.*, 2005). The resulting GO terms were further generalized into the set of terms relevant for plant taxa using the reduced GO-Slim plant set (http://www.geneontology.org/).

Differential expression. Reads from each of the 4 root-tip libraries were aligned back to the *de novo* transcriptome using TopHat, reporting only the best alignment for each read. The number of reads aligned per contig were tabulated using Perl scripts, and used as inputs to the program

DESeq (Anders & Huber, 2010) for calculating the significance of fold changes in expression between the control and Al treatments for each genotype. Read counts from each library represent an average value of biological replicates (which were pooled during tissue collection). Confidence intervals for gene expression levels are thus based on variation inherent in sampling read data.

Quantitative PCR. Twenty-four unigenes were selected from the larger data set for quantitative (qPCR) validation and further analysis. The unigenes we selected had the following characteristics: >1 fragment per kilobase per million base pairs (FPKM) coverage in the *de novo* assembly, >10 reads per kilobase representation in each separate treatment library, significant regulation in response to Al in one or both genotypes, and expression that differed in magnitude between the sensitive and tolerant genotype. Most of the tested transcripts had putative function related to Al tolerance mechanisms while some were undescribed (marked in Table 6 and 7). Expression of these genes was measured in control and Al treated root tips from 7 PGE plant genotypes that ranged widely in Al tolerance (Fig. 1). Three to five replicate tillers from each genotype were control and Al treated, root tip tissue was collected, and RNA extracted as described above for the RNA-seq plants. Root tip collection was made 24 hours after exposure to Al. cDNA libraries were prepared for each sample using iScript cDNA synthesis **(**BioRad).

We designed qPCR primers to amplify a portion of each unigene sequence (ABI Primer Express 3.0). When highly similar transcripts were detected in the assembly, we manually designed primers to amplify sequence portions that had the highest number of differences between them. Also, in some cases we detected groups of related alternative splice forms based on Trinity clustering and hierarchical clustering alignments (MultAlign, (Corpet, 1988)) and so

designed primers to detect only the form(s) that showed evidence of differential regulation in the RNA-seq dataset. qPCR analysis was performed in 384-well plates using SYBR Green reagents on the ABI platform. *Histone-3* was used as a control gene for the assay and expression was quantified using the –ddCt method (Livak & Schmittgen, 2001) with 2 to 5 replicates per unigene sequence per genotype per treatment. Significance of gene expression differences between control and Al exposed tissue within each genotype was calculated using t-tests.

Results

Tissue Al content. We treated *Anthoxanthum* seedlings with Al over a period of 4 days and measured uptake of into leaves, root cell walls, and root symplast (sap). We found Al was clearly partitioned between plant tissues (Fig. 2a). On average, leaf tissue had much less Al than either root tip cell walls or sap, and the Al content of different tissues was not correlated across samples (Table 2). There was a high degree of variation in cell sap Al content between individuals but no strong correlation between that and tolerance (Fig. 2b). Plants with higher Al content in the root cell wall were, however, less tolerant than those that excluded Al (Fig. 2c). There were generally weak negative correlations between the nutrient content of leaves and the Al content of both root cell walls and sap, but RRG (tolerance) was not significantly associated with higher nutrient contents, possibly with the exception of iron (corr = 0.49).

The *Anthoxanthum* root tip transcriptome. Following sequencing and assembly, the final *de novo* root tip transcriptome contained 86,488 unigenes reconstructed from across two tetraploid

genomes. While most unigenes in the set are likely to be transcribed from separate genomic loci, some are also likely to be divergent alleles at the same locus either within or between genomes. Some unigenes also represent sets of related alternative splice forms, which was evident in hierarchical clustering alignments of similar sequences (data not shown). The assembly had high coverage per unigene with an average of 15 FPKM. The mean unigene length was 754 bp, and half of all base pairs were incorporated into unigenes 850 bp in length or greater (Table 3).

We were able to describe a high percentage of sequences in the transcriptome using comparative information from cultivated grasses and other plants (Fig. 3). 52.8% of unigenes had significant protein similarity to previously described plant genes (at E-value $<10^{-6}$), and overall the transcriptome had more similarity to more closely related grasses (*Brachypodium,* barley, wheat, sorghum, rice, and corn respectively, Fig. 3 and Table 4). Sixty-two percent of the unigenes that had significant BLAST hits could be reliably assigned functional GO terms (32.8% of all unigenes). The annotated unigenes represented 17 diverse categories of molecular functionality (Fig. 3A). Of note, a large proportion of sequences had nucleotide, DNA, or protein binding capacity, kinase activity and transporter activity.

Functional term enrichment analysis. Expression of more than 91% of all unigenes in the transcriptome could be reliably measured in each root tip library (Table 1). We grouped those unigenes into categories based on their expression patterns (Table 5). The experiment had a low amount of noise as evidenced by the fact that 98.2% of unigenes showed no significant expression change between control and Al treatments in either plant. Only 133 unigenes (0.01%) were significantly up- or down-regulated in both genotypes in response to Al. Sixty-seven of these could be assigned GO functionality. When compared with the set of all functionally

annotated unigenes in the transcriptome (n = 28,354), these broadly Al responsive unigenes were enriched for extra-cellular region proteins and for proteins involved in biosynthesis, gene expression, translation, DNA and nucleotide binding, catalysis, and metabolism of nitrogen-containing compounds (Fig. 4A).

Because differences in gene expression have the potential to underlie adaptive differences in Al tolerance at PGE, we also looked more closely at unigenes with significant expression differences between the tolerant and sensitive plant genotypes. Regarding constitutive differences (in the absence of Al exposure) 2,902 unigenes had significant root expression differences between the two plant genotypes and 615 of them could be assigned functionality. Among them function was enriched for kinase activity, cell death, response to stress, epigenetic gene regulation and chromatin binding (Fig. 4B). We also identified a set of 124 unigenes with a large difference in response to Al between the two genotypes (defined as a difference greater than 0.5 in the Al responsive expression fold-change between genotypes, regardless of the direction of expression changes). Eighty-five of these sequences could be assigned GO functionality and among them function was enriched for metabolism, biosynthesis, and catalysis, and, uniquely, lipid metabolism (Fig. 4C).

Al responsive transcripts in *Anthoxanthum.* In order to more specifically identify transcripts potentially involved in Al tolerance differences between plant genotypes we looked at BLAST descriptions of unigenes with expression differences that have either >2x or <0.3x fold-change under Al treatment.in the tolerant plant genotype (but many of them also had the same pattern in the sensitive genotype). Among them we found 33 unigenes with descriptions similar to Al response transcripts identified in studies of other plant species (Houde & Diallo, 2008; Kumari *et*

al., 2008; Chandran *et al.*, 2008; Sawaki *et al.*, 2009; Yamaji *et al.*, 2009; Grisel *et al.*, 2010; Delhaize *et al.*, 2012) (Table 7). There were also 121 unigenes in the group with no significant similarity to previously described sequences.

This set of Al responsive unigenes includes several sequences involved in transport of ions across membranes. Three metal ion-specific *Nramp* transporters, two magnesium transporters, one proton antiporter, and three ATP-powered transporters (atp-binding cassette proteins) were up-regulated in one or both genotypes in response to Al (Table 6). Transcripts with similarity to genes controlling organic acid efflux from root tips were also up-regulated including one with similarity to ALMT1, an important malate efflux protein linked to Al tolerance variation in wheat and barley (Sasaki *et al.*, 2004). Also of note, some MATE-family unigenes were regulated in response to Al in *Anthoxanthum*, which confer tolerance in sorghum, corn and others (Magalhaes *et al.*, 2007; Furukawa *et al.*, 2007; Maron *et al.*, 2010; Yokosho *et al.*, 2011).

Several transcripts putatively involved in cell wall modification were also up-regulated in response to Al. These include glucosyltransferases, which transfer sugars during cell-wall remodeling and growth (Schopfer, 2001; Liszkay *et al.*, 2004), and unigenes with similarity to enzymes that modify cell-wall components such as hemicellulose (Nishitani & Tominaga, 1992; Zhu *et al.*, 2012). As was evident in the GO term enrichment analysis, many unigenes involved in biosynthesis and metabolism were also differentially regulated. These included unigenes associated with translation and protein synthesis, such as transfer RNAs and their synthases, as well as ribosomal proteins and one cytochrome p450. Several unigenes putatively involved in synthesis of terpenes (secologanin synthases) were also up-regulated, but the significance of this is unknown.

Interspecies comparison of Al response genes. We conducted direct BLAST searches of the *Anthoxanthum* transcriptome for homologs of 16 major Al tolerance genes that have been identified in other plant species (Table 7). All but two loci had high similarity matches (alignment scores > 100, Table 7). The organic acid exporters ZmMATE (citrate transporter) and TaALMT1 (malate transport), as well as the vacuolar Al transporter OsNrat1 and the ABC-transporter AtALS1 all have putative homologs that are up-regulated in response to Al in *Anthoxanthum*. Counter to expectation, up-regulation of the two organic acid transporters was greater in the sensitive than the tolerant plant genotype. However, we did observe that constitutive expression for some of the putative homologs was slightly higher in the tolerant genotype (e.g. ZmMATE and OsNrat1). The majority of other tolerance loci with putative homologs in *A. odoratum* were highly expressed in both control and Al exposure treatments but showed no evidence of differential regulation in response to Al.

Relationship between candidate gene expression level and Al tolerance. We tested the expression patterns of a subset of candidate unigenes (n = 22, Fig. 5) to examine the relationship between Al tolerance and expression across genotypes with different levels of Al tolerance (n=7). In general when up-regulation of unigene expression was predicted based on RNA-sequencing, it was confirmed across most genotypes (Fig. 5). Interestingly, the level of expression for these unigenes tended to be negatively correlated with Al tolerance (RRG, Table 8). The strength of correlation varied (Table 8, R^2 values) but was often quite strong. We detected no correlation between constitutive expression of any of these unigenes and RRG.

Discussion

The ability to thrive in Al-toxic soils is a trait that has been artificially selected in cereal crops but is also a naturally evolved trait in some wild members of the Poaceae. Here we present the first comparative analysis of the genetic basis of soil Al tolerance between cereal crops and a wild grass: *Anthoxanthum odoratum* (sweet vernal grass). We looked broadly at the Al content of tissues across genotypes of different tolerance and then used RNA sequencing to characterize the transcriptome of root tips, the primary site of Al stress in acid soils. We compared Al responsive transcripts with loci known to be important in cultivated grasses to address whether Al tolerance has a similar or different genetic basis across species. Lastly, we looked at the relationship between Al tolerance and gene expression to determine if gene expression differences may contribute to locally adaptive differences within the species.

Transcriptomic response to Al. Broadly we found that both metabolism and biosynthesis are regulated in *Anthoxanthum* in response to Al exposure. Up-regulation of genes in these pathways may be a precursor of induced biotic and abiotic stress responses (Mittler, 2002). Translation and biosynthesis are activated to produce protective or defensive compounds. Metabolism is up-regulated in order to provide energy and building blocks, and sometimes later to repair or regenerate damaged cellular components. It is likely that *Anthoxanthum* has evolved one or even many induced responses to Al stress, which may contribute to its remarkable tolerance. The number of unigenes that are differentially regulated in roots in response to Al is on par with that found in other species over short (=<48 hour) exposure periods (at approximately 2x up- or down- regulation). We identified a total of 267 Al regulated unigenes in *Anthoxanthum* roots

while similar studies have identified 450 in maize (Maron *et al.*, 2008), 264 in poplar (Grisel *et al.*, 2010), 1,114 in *Arabidopsis* (Kumari *et al.*, 2008), and 591 in *Medicago* (Chandran *et al.*, 2008). Many of the genes responsive to Al stress are similar across species and we find the same is true for *Anthoxanthum*. We also see that a large number of transcripts of unknown function (121) are significantly regulated in response to Al, and it is likely that most are involved in generating Al tolerance of *Anthoxanthum*. In particular the un-annotated unigene UN71256 showed both a high level of up-regulation (90-135 fold increase, Table 6) as well as large variation in expression between genotypes (Fig. 4). In addition, some unigenes we identified have protein similarity to other grasses, but their function in Al stress response may be novel. For example, several unigenes with similarity to secologanin synthases were identified as regulated in response to Al, but this enzyme is involved in alkaloid biosynthesis, which has thus far not been linked to Al stress response or tolerance in plants. It is also possible that changes in regulation at the level of transcripts are not the primary cause of Al tolerance in *Anthoxanthum.* Preliminary analysis shows that hundreds to thousands of unigenes are constitutively expressed to a much greater degree in the tolerant plant versus sensitive plant at the adult stage. Physiological testing combined with further expression comparisons across genotypes and growth stages are required to investigate the significance of such constitutive expression differences.

Interspecies comparisons and the mechanism of Al tolerance. Based on interspecies comparisons, the evidence here suggests that two major types of Al tolerance mechanisms are likely to be active in *Anthoxanthum*: external Al exclusion and internal detoxification (Ma, 2007; Delhaize *et al.*, 2012). Exclusion of Al via excretion of organic acids at the root tip is suggested

by the presence and up-regulation of major organic acid excretion loci for both malate (ALMT1) and citrate (MATE). We detected at least 6 transcripts with high similarity to TaALMT1 in *Anthoxanthum* (data not shown), and two of them, UN14079 (Table 6) and UN11139 (Table 7, Fig. 4), were up-regulated in response to Al in sensitive and tolerant plants. Similarly, several transcripts with high similarity to the homologous citrate excretion loci ZmMATE1 and OsFRDL4 were identified in *Anthoxanthum*, two of which (UN11622 and UN08191) were up-regulated in the Al treatments.

Despite the apparent regulation of these transcripts, the overall significance of root tip exclusion of Al in *Anthoxanthum* remains unclear. Although Al is low in the leaves, we find Al does enter root cells in moderate amounts (Fig. 1). In preliminary tests using simple paper chromatography we have not detected exudation of malate, citrate, or oxalate from Al exposed root tips (data not shown). It is possible that organic acid excretion occurs only at a very low level in this plant and is not the primary mechanism of dealing with Al toxicity, as also appears to be the case for rice, the most tolerant of cultivated grasses (Famoso *et al.*, 2011). It is also possible that in *Anthoxanthum* organic acid production functions primarily in internal Al chelation and transport (see (Pineros *et al.*, 2008)). Targeted functional testing is required to test these hypotheses.

Cell-wall modification is another Al exclusion mechanism that may be active in *Anthoxanthum,* although not a primary mechanism of tolerance. Some of the most well-understood loci involved in Al induced cell-wall modification are OsSTAR1 and OsSTAR2 which complex in order to transport UDP-glucose to the cell wall, possibly in order to block damage by Al^{3+} ions (Huang *et al.*, 2009). Unigenes with high similarity to both OsSTAR1 and OsSTAR2 were identified in *Anthoxanthum* and are both (weakly) up-regulated in response to

Al, as are several other transcripts potentially involved in sugar transport. Sugar transferases have been shown to be important in Al tolerance in rice (Yang *et al.*, 2008). However, significant amounts of Al are still present in *Anthoxanthum* cell walls indicating Al exclusion is weak, at least under these experimental conditions.

Internal detoxification and sequestration of Al may also be important in *Anthoxanthum*. Leaf sequestration is the most well understood mechanism and is known only in non-grass plants such as buckwheat, hydrangea, and tea (Taylor *et al.*, 1997; Wenzl *et al.*, 2001; Metali *et al.*, 2011). In these species Al is immediately chelated and transported to the leaf tissues where it is stored in the vacuole, reaching concentrations of 1000's of micrograms of Al per gram of leaf tissue. However, based on tissue Al content, it is unlikely that *Anthoxanthum* uses Al leaf tissue sequestration as a mechanism of Al tolerance even though it is far more Al tolerant than most non-accumulating herbaceous plants so far tested. In fact, it is remarkable that at 950 uM direct hydroponic exposure, the level of Al in *Anthoxanthum* seedling leaf tissue remains undetectable even after 4 days of growth (Fig. 1). Where accumulation of Al is evident to some degree is in the root tissues, both in cell walls and inside the cells. The only other plant where sequestration in roots has been detected to some extent is rice (Xia *et al.*, 2010; Huang *et al.*, 2011). Unigenes with high similarity to plasma membrane Al^{3+} ABC-transporter OsALS1 and Al-specific *Nramp* protein involved in vacuolar sequestration, OsNrat1, are present in *Anthoxanthum* and weakly up-regulated. In addition, at least three *Nramp* proteins and 3 other ABC-transport proteins were identified, although their status as homologs or homeologs (differentiated members of one tetraploid locus) is unknown. This taken together suggests that internal transport, detoxification, and sequestration of Al may be an Al stress response in *Anthoxanthum*.

Intraspecies variation of gene expression and the evolution of local adaptation. The role of transcript expression differences in the evolution of local adaptation to soil Al remains unclear from the results observed here. We were able to detect significant differences between individuals of different Al tolerance levels in the expression of putative homologs of Al tolerance genes in crops, however expression was not correlated with tolerance. We measured the relationship between Al tolerance and candidate gene expression in 22 targets, and counter to expectation most were more highly expressed in Al sensitive rather than Al tolerant plants. This might be explained by lack of precision in the measure of Al tolerance (RRG), which varies between replicate measures of the same individual. Also, if the correlation between gene expression and tolerance is weak, a panel of hundreds of plants rather than the smaller set used here, might be required to detect a correlation. However, the consistency of the negative relationship between Al tolerance homologs and expression level may also suggest a third alternative: Variation for Al tolerance in *Anthoxanthum* may be linked to a novel mechanism of Al exclusion, and expression of most other Al response genes occurs secondarily in response to Al damage. In this scenario we would predict that plants with higher Al tolerance would sustain less damage by Al and thus have a lower level of secondary up-regulation of canonical Al response loci. The initial, primary tolerance mechanism could be linked to one of the response loci we did not measure across genotypes, or one or more of the 121 loci that are regulated in response to Al but have no significant similarity to known sequences. These provide exciting targets for future research.

Tools such as RNA sequencing provide an excellent starting point for exploring the diversity of genetic evolution across organisms that do not have extensive pre-existing genetic resources, like *Anthoxanthum*. Physiological and functional genetic assays however are then

required to decipher the role of the genes that are uncovered. For Al tolerance and other ecologically important traits, this next step should be an important component of future research programs given what little information exists on the genetic basis of functional biodiversity outside of the small phylogenetic representation of current model species.

Acknowledgements: Special thanks for help and support from John Storkey, Eric Craft, Jon Shaff, Adam Famoso, Chih-Wei Tung, Jianyong Li, Zhangjun Fei, Julie Oh, and members of the Cornell Computational Biology Service Unit. Funding for this work was provided by a grant from the Cornell Center for Comparative Genomics, the Cornell chapter of Sigma Xi, and the Biogeochemistry IGERT.

FIGURES

Figure 1. Al Tolerance of plant Genotypes from PGE used for gene expression measurements. Grey points are genotypes used for qPCR measurements; white point is the sensitive RNA-Seq genotype; black point is the tolerant RNA-seq genotype. Bars represent +/- 1S.E.

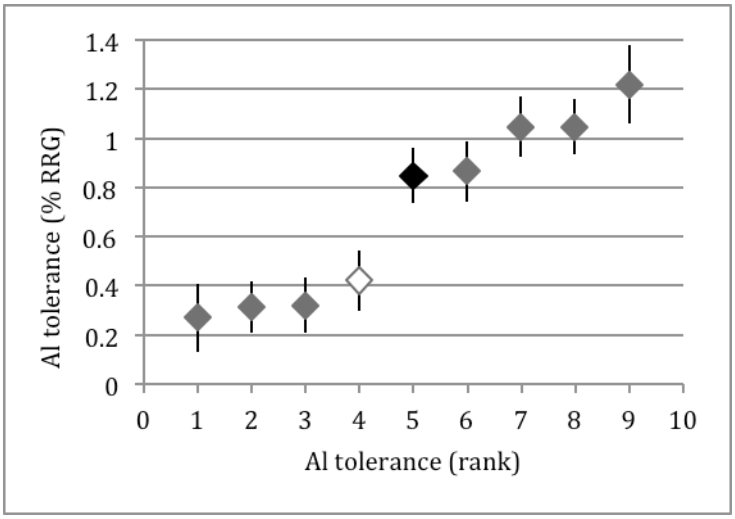
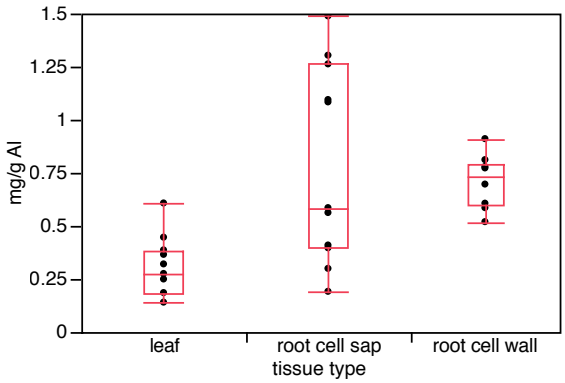
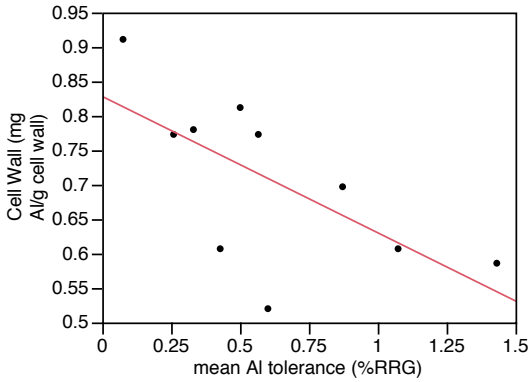


Figure 2 a) Al content of *Anthoxanthum* tissues and b-c) correlation with Al tolerance (relative root growth, RRG). Tissue was collected from 11 bulked tissue samples (see methods) of seedlings that were exposed to 950 uM Al for four days. *Cell sap Al content was standardized by mg cell wall in the original sample.

a) Al Content of *Anthoxanthum* tissues



b) Cell wall (p=0.04)



c) Cell Sap* (p=0.68)

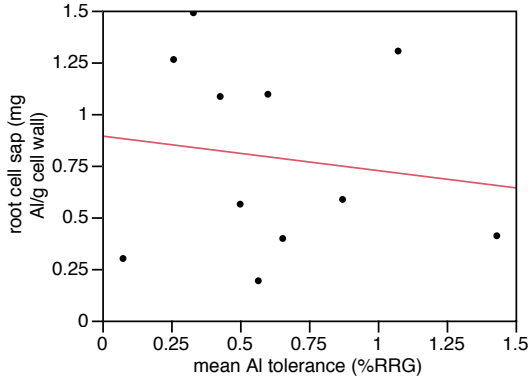
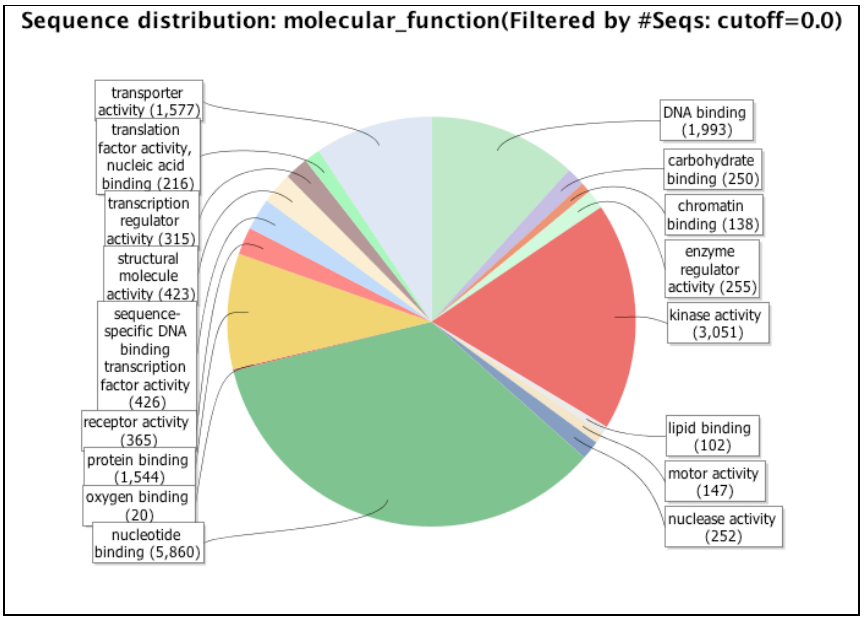
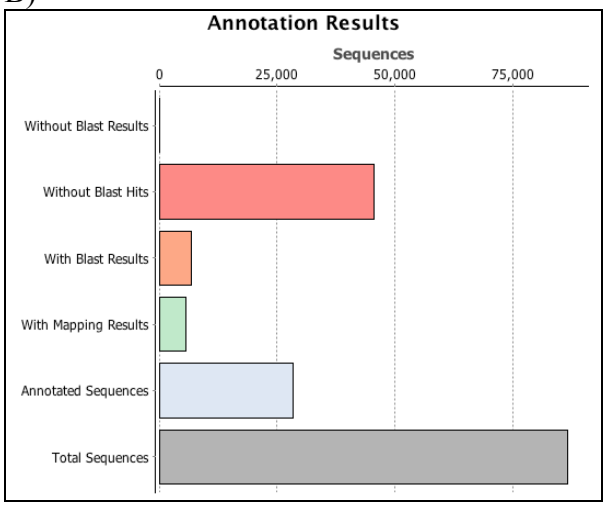


Figure 3. Whole root tip transcriptome annotation statistics. A) percentage of GO functionalities among all annotated sequences. B) number of sequences passing each stage in the Blast2Go annotation pipeline. C) number of sequences with top scoring BLAST hits to plant species in GenBank.

A)



B)



C)

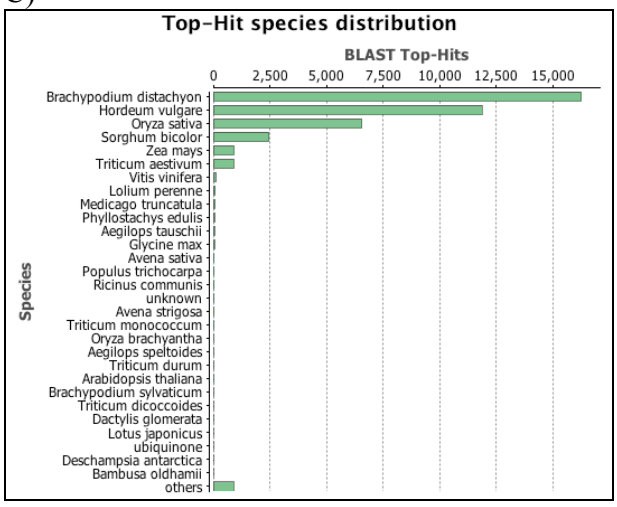
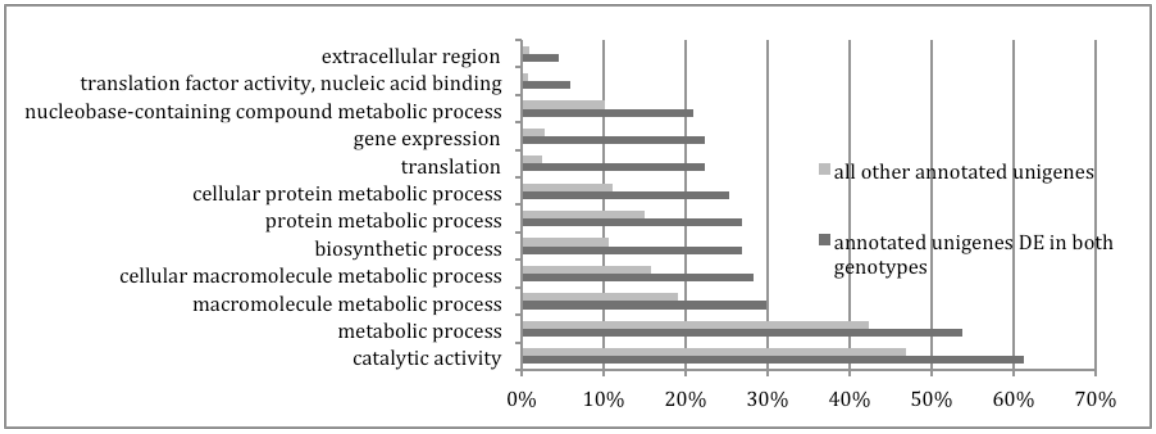
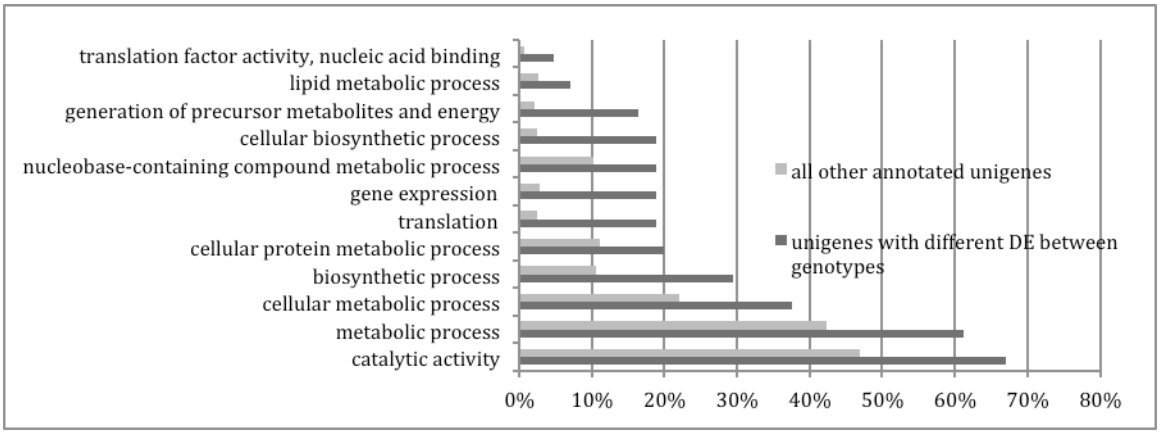


Figure 4. GO Term Enrichment Analysis**.** Results are shown for GO categories overrepresented at $p<0.05$, Fisher's Exact test, among A) unigenes that show differential expression (DE) in response to Al exposure in both the sensitive and tolerant plant genotypes (n=67); B) unigenes with a >0.5 fold-change difference in the amount of DE between genotypes (n=85); and C) unigenes with significant constitutive expression differences between plant genotypes (n= 615). Enrichment is compared with a reference set of all unigenes that can be assigned GO functionalities (n=28,288). Note, some GO terms are overlapping.

A)



B)



C)

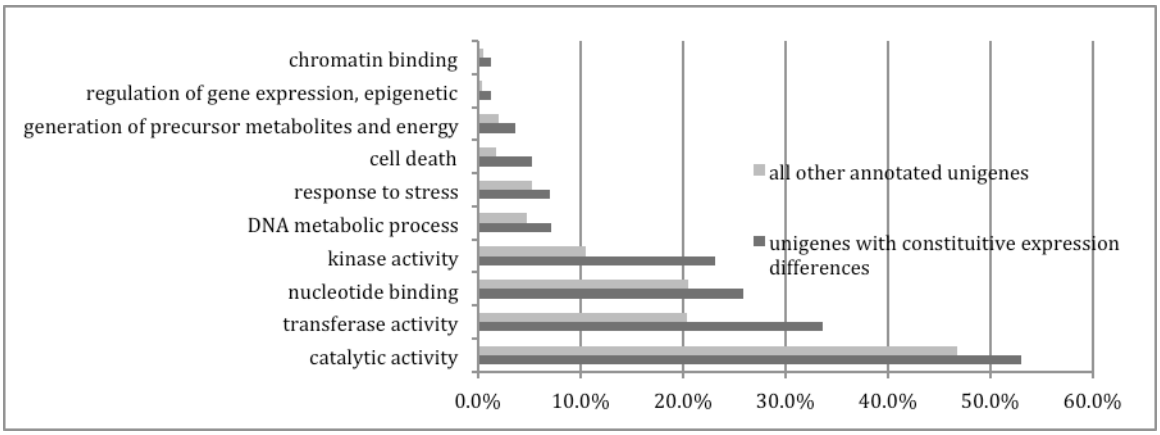
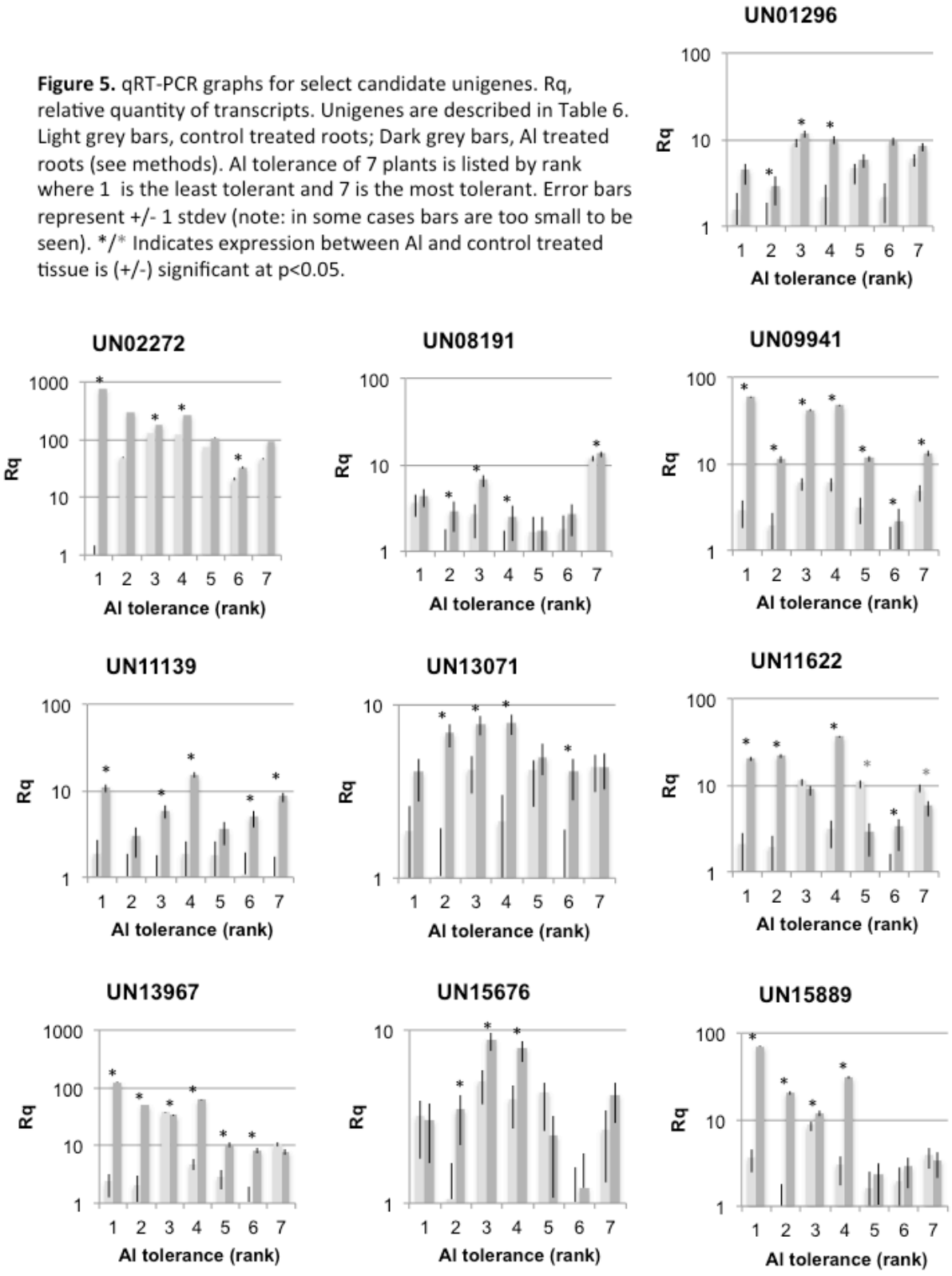
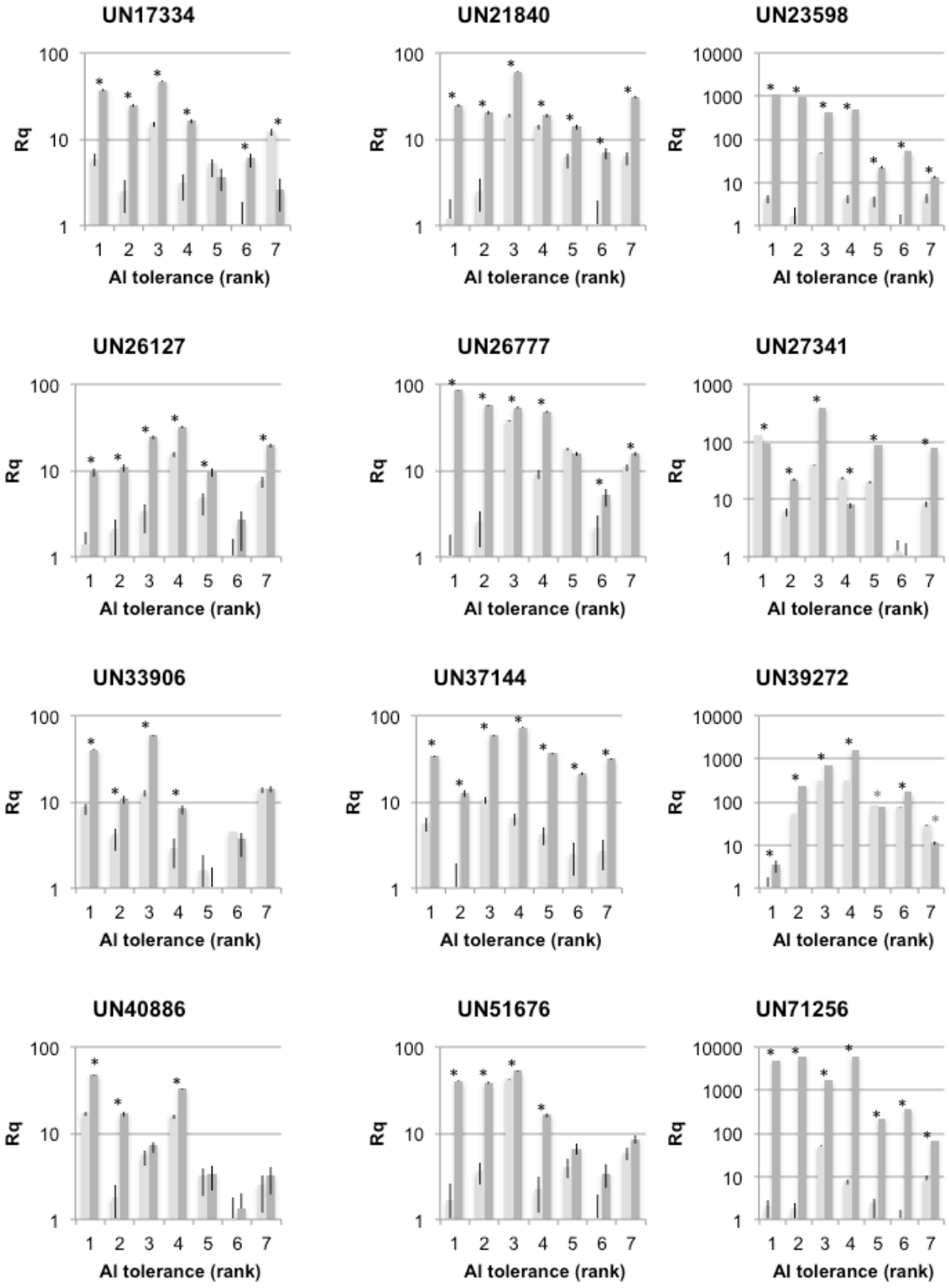


Figure 5. qRT-PCR graphs for select candidate unigenes. Rq, relative quantity of transcripts. Unigenes are described in Table 6. Light grey bars, control treated roots; Dark grey bars, Al treated roots (see methods). Al tolerance of 7 plants is listed by rank where 1 is the least tolerant and 7 is the most tolerant. Error bars represent +/- 1 stdev (note: in some cases bars are too small to be seen). */* Indicates expression between Al and control treated tissue is (+/-) significant at $p<0.05$.




UN17334
UN21840
UN23598
UN26127
UN26777
UN27341
UN33906
UN37144
UN39272
UN40886
UN51676
UN71256
Rq
Al tolerance (rank)

TABLES

Table 1. RNA-Seq Alignment Statistics for 4 libraries aligned to the combined *de novo* transcriptome (unigene set).

Library No.	Plant Phenotype	Treatment	Total reads ($x10^6$)	# Reads, filtered ($x10^6$)	Avg base quality score* (PHRED)	% reads aligned to reference	% Unigenes hit (>=10 RPK)	average unigene expression (RPKM)
1	sensitive	control	155.8	103.6	33.2	67%	91.2%	10.4
2		+Al	184.0	143.9	33.4	78%	91.1%	10.4
3	tolerant	control	152.3	128.3	33.4	84%	91.2%	10.3
4		+Al	210.5	106.0	33.4	50%	91.3%	10.3

Table 2. Pearson correlations between Al tolerance, Al content in tissues and nutrient content in leaves of seedlings. N = 11 pooled tissue samples (see methods). Strong correlations are in bold. cw: root cell wall, sap: root cell symplast.

	Al tolerance (%RRG)	sap Al	cw Al	leaf Al	Ca	Mg	K	P	B	Mn	Fe	Cu	Mo
sap Al	-0.14												
cw Al	**-0.65**	-0.33											
leaf Al	0.22	-0.08	-0.44										
Ca	0.32	-0.38	-0.03	0.22									
Mg	0.40	-0.26	-0.16	0.39	**0.93**								
K	-0.02	-0.42	-0.19	**0.65**	0.00	0.08							
P	0.11	0.12	-0.46	0.36	0.42	**0.51**	0.40						
B	0.24	0.19	**-0.69**	0.24	0.13	0.06	0.30	**0.53**					
Mn	0.33	-0.14	0.03	0.13	**0.81**	**0.83**	-0.02	0.33	0.17				
Fe	-0.49	-0.28	**0.57**	-0.09	**0.56**	0.40	0.01	0.20	-0.02	0.37			
Cu	-0.13	-0.24	-0.04	0.28	0.46	**0.54**	0.48	**0.85**	0.19	0.29	0.48		
Mo	0.38	-0.05	-0.10	0.18	**0.69**	**0.83**	-0.12	0.24	-0.23	**0.79**	0.12	0.23	
Zn	-0.32	0.06	-0.06	0.60	0.47	0.50	0.35	0.61	0.34	0.34	0.48	0.61	0.20

Table 3 Transcriptome assembly statistics.

Input reads	% Reads Assembled	# Unigenes	Avg. FPKM*	Mean Unigene length	Median Unigene length	N50
96M	47%	86,488	15.0	754 bp	547 bp	824 bp

Table. 4 *Anthoxanthum* sequence similarity to model plant species. Similarity is calculated as the number of top BLAST hits to a species divided by the number of sequences available in the nr database for that species.

Species	**# sequences in nr database**	**# top BLAST hits in *Anthoxanthum***	**Similarity (% available seqs hit)**
Brachypodium distachyon (brome)	25,545	16,256	64%
Hordeum vulgare (barley)	34,685	11,876	34%
Triticum aestivum (bread wheat)	10,748	910	8%
Sorghum bicolor (sorghum)	71,876	2,452	3%
Oryza sativa (rice)	280,612	6,562	2%
Zea mays (corn)	108,085	915	1%
Arabidopsis thaliana (thale cress)	224,411	30	<0.001%

Table 5. Number of unigenes regulated in response to 24 hour Al exposure in one sensitive and one tolerant plant genotype. Up- and down- regulated unigenes have significant expression differences between the control and Al treatments (within a genotype) at $p<0.05$ (see Methods).

		tolerant genotype		
	regulation	up	no change	down
sensitive genotype	up	94	274	0
	no change	84	84,971	40
	down	10	976	39

Table 6. Selected differentially regulated unigenes. Unigenes shown in the table have putative function related to known Al tolerance mechanisms and/or significant up or down regulation in either both or only the tolerant plant genotype under Al exposure. –Al, control treatment; +Al, Al treatment; BaseMean value = #reads/ library geometric mean #reads. Os, *Oryza sativa*; At, *Arabidopsis thaliana*; Zm, *Zea mays*; Ta, *Triticum aestivum*. Bold values indicate up or down regulation is statistically significant (DESeq adjusted p-value <0.05). + transcripts measured across genotypes using qPCR (see Figure 5). *an additional 119 unigenes of unknown function are not shown here.

Candidate Unigene	Description	length (bp)	Al Sensitive (BaseMean)		Al Tolerant (BaseMean)		Sen fold change	Tol fold change
			-Al	+Al	-Al	+Al		
Transport								
UN11139+	ALMT1	478	148	2,014	297	1,456	**13.6**	**4.9**
UN40932	magnesium transporter mrs2-e-like	640	494	1,568	286	1,324	**3.2**	**4.6**
UN03316	magnesium transporter mrs2-e-like	540	2,108	4,319	969	3,243	2.0	3.3
UN13369	cation h(+) antiporter 15-like	1382	22	869	17	407	**40.1**	**24.5**
UN01296+	metal transporter	505	188	387	139	865	2.1	**6.2**
UN37144+	nramp transporter	586	4,283	15,044	3,102	7,875	**3.5**	2.5
UN33906+	nramp transporter	385	66	22	17	51	0.3	3.1
UN11622+	MATE efflux family protein 5-like	1078	2,226	4,444	1,507	4,182	2.0	2.8
UN40886+	abc transporter-like	464	5	48	1	357	10.6	**365.4**
UN26777+	abc transporter e family member 2-like	377	554	903	195	2,014	1.6	**10.3**
UN04216	multidrug resistance protein abc transporter family	576	669	1,414	260	1,451	2.1	**5.6**
Cell Wall Modification								
UN10052	udp-glucosyltransferase	678	357	1,770	47	685	**5.0**	**14.6**
UN51676+	utp-glycosyltransferase-like	1213	24	91	22	125	3.8	5.6
UN66043	beta glucanase	634	1,220	109	895	76	**0.1**	**0.1**
UN17990	endoxyloglucan transferase	386	388	145	949	205	0.4	**0.2**
UN15676+	subtilisin-like protease-like	763	1,634	1,498	1,180	2,248	0.9	1.9

Table 6 *continued*

Candidate Unigene	Description	length (bp)	Al Sensitive (BaseMean)		Al Tolerant (BaseMean)		Sen fold change	Tol fold change
			-Al	+Al	-Al	+Al		
Metabolism								
UN09941+	cytochrome p450	1304	1989	2428	212	737	1.2	3.5
UN13967+	methionyl-trna expressed	1064	806	5498	703	6030	**6.8**	**8.6**
UN17334+	bifunctional aminoacyl-trna expressed	664	52	384	64	2006	7.3	**31.1**
UN15889+	eukaryotic translation initiation factor 6+	352	128	1928	94	2116	**15.1**	**22.6**
Other functions								
UN02272+	transcription factor	639	316	1,377	239	2,572	**4.4**	**10.8**
UN02544	root abundant factor (transcription factor)	614	315	35	360	50	**0.1**	**0.1**
UN19546	root abundant factor (transcription factor)	576	486	89	166	11	**0.2**	**0.1**
UN49366	auxin response factor	397	1	11	9	313	9.2	**35.6**
UN55833	auxin response factor 17-1	374	3	15	2	209	4.4	**107.1**
UN51563	auxin response factor 13-like	524	1	2	7	364	1.5	**53.2**
UN34546	blue copper binding protein	692	38	94	12	132	2.5	11.3
UN27341+	heavy metal-associated domain-containing protein	1026	310	86	84	435	0.3	5.2
UN10824	secologanin synthase-like	1649	2,084	2,261	678	2,918	1.1	**4.3**
UN21840+	secologanin synthase-like	1570	271	499	102	1,054	1.8	**10.3**
UN23598	secologanin synthase-like+	426	78	695	20	875	**9.0**	**42.7**
UN26127+	secologanin synthase-like	1641	132	254	29	636	1.9	**21.7**
UN53162/3	secologanin synthase-like	540	928	1,369	88	623	1.5	**7.1**
UN71256+	unknown*	370	22	1,968	23	3,168	**90.8**	**135.3**
UN39272+	unknown*	442	30	27	1,617	8,654	0.9	**5.4**

Table 7. Best *Anthoxanthum* matches to Al tolerance loci from model plant species. [+]transcript measured across genotypes using qPCR (see Figure 5).

Gene(s)	Putative *Anthoxanthum* homolog	BLAST score	E-value	unigene length (bp)	BLAST Description	Al Sensitive expression		Al Tolerant expression		Sen fold change	Tol fold change
						N	A	N	A		
OsALMT1	UN14079	381	2.0E-128	1120	ALMT1	794	8,882	626	3,351	**11.2**	**5.4**
TaALMT1		456	1.0E-157								
OsFRDL4	UN08191[+]	304	6.0E-96	1279	citrate efflux mate transporter	12,650	35,397	13,594	37,354	**2.8**	2.7
ZmMATE		492	2.0E-169								
OsNrat1	UN08554/5	282	3E-91	882	metal transporter nrat1	59,311	277,473	65,284	166,925	**4.7**	2.6
OsSTAR1	UN11920	243	2E-77	1041	phosphate import atp-binding protein pstb 1	3,016	9,656	2,064	5,885	**3.2**	2.9
OsSTAR2	UN50094/5/6	283	2E-92	1123	protein aluminum sensitive 3	3,002	6,199	112	115	2.1	1.0
AtALS3		260	7E-85								
AtALS1	UN13071[+]	865	0.00	888	abc transporter	29,867	70,247	24,741	39,888	**2.4**	1.6
OsART1	UN65068	209	1E-60	1528	c2-h2 zinc finger protein	6,215	5,562	4,845	4,473	0.9	0.9
OsMGT1	UN33791	219	5E-64	1821	retrotransposon ty3-gypsy subclass	9,355	13,719	9,887	14,720	1.5	1.5
OsISL	UN82594	120	2E-31	386	isocitrate lyase	2	2	31	70	0.8	2.2
AtATR	UN01856	545	2E-174	692	serine threonine-protein kinase atr-like	789	1,145	1,021	931	1.5	0.9
AtALT2	UN77886	59.3	2E-09	1214	katanin p80 wd40 repeat-containing subunit b1-like	2,879	2,998	3,409	2,966	1.0	0.9
AtSTOP1	UN65068	347	2E-112	1528	c2-h2 zinc finger protein	6,215	5,562	4,845	4,473	0.9	0.9
AtBCB	UN84836	80.9	8E-19	409	predicted protein	1,041	1,565	483	983	1.5	2.0

Table 8. Regression analysis of qPCR data. Expression was measured for 22 unigenes that showed a different magnitude of differential expression in the sensitve and tolerant RNA-s genotypes. Expression was measured across 7 plants that rank in Al tolerance from very se to very tolerant (1 to 7, respectively, see Fig. 1). We plotted a linear regression for the relationship between transcript expression and tolerance for each locus, in both control and exposure treatments. The slope and R^2 values of each regression are shown below.

	+Al		Control	
Unigene	**regression slope**	**R^2**	**regression slope**	**R^2**
UN23598	-197.82	0.88	-1.60	0.04
UN26777	-12.65	0.88	0.31	0.00
UN13967	-16.56	0.70	-0.46	0.01
UN51676	-7.62	0.69	-1.12	0.03
UN02272	-92.86	0.67	0.83	0.00
UN17334	-6.68	0.67	0.22	0.01
UN15889	-8.80	0.60	-0.16	0.02
UN71256	-989.70	0.60	-1.01	0.01
UN40886	-6.03	0.53	-1.70	0.29
UN09941	-6.64	0.42	0.03	0.00
UN11622	-3.12	0.29	0.72	0.11
UN33906	-5.44	0.29	0.18	0.01
UN01296	0.71	0.22	0.41	0.09
UN08191	0.78	0.17	0.92	0.26
UN13071	-0.27	0.11	0.27	0.14
UN21840	-1.98	0.06	-0.03	0.00
UN27341	-14.34	0.05	-13.92	0.46
UN15676	-0.26	0.04	-0.09	0.01
UN11139	-0.17	0.01	-0.08	0.13
UN39272	-27.33	0.01	-3.80	0.00
UN26127	-0.05	0.00	0.65	0.07
UN37144	-0.45	0.00	-0.44	0.09

Literature Cited

Anders S, Huber W. **2010**. Differential expression analysis for sequence count data. *Genome Biology*. **11**: 1–12.

Antonovics J. **1972**. Population dynamics of the grass *Anthoxanthum odoratum* on a zinc mine. *Journal of Ecology* **60**: 351–365.

Arendt J, Reznick D. **2008**. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends in Ecology & Evolution* **23**: 26–32.

Chandran D, Sharopova N, Ivashuta S, Gantt J, Vandenbosch K, Samac D. **2008**. Transcriptome profiling identified novel genes associated with aluminum toxicity, resistance and tolerance in Medicago truncatula. *Planta* **228**: 151–66.

Chen ZC, Yamaji N, Motoyama R, Nagamura Y, Ma JF. **2012**. Up-regulation of a magnesium transporter gene OsMGT1 is required for conferring aluminum tolerance in rice. *Plant Physiology* **159**: 1624–33.

Conesa A, Götz S, García-gómez JM, Terol J, Talón M, Genómica D, Valenciano I, Agrarias DI, Valencia UP De. **2005**. Blast2GO : a universal tool for annotation , visualization and analysis in functional genomics research. **21**: 3674–3676.

Corpet F. **1988**. Multiple sequence alignment with hierarchical clustering. *Nucl. Acids Res.* **16**: 10881–10890.

Davies MS, Snaydon RW. **1973**. Physiological differences among populations of *Anthoxanthum odoratum* L. collected from the Park Grass Experiment, Rothamsted. II. Response to Aluminium. *Journal of Applied Ecology* **10**: 47–55.

Davies MS, Snaydon RW. **1976**. Rapid population differentiation in a mosaic environment. III. Measures of selection pressures. Journal name is missing **36**: 59–66.

Delhaize E, Ma JF, Ryan PR. **2012**. Transcriptional regulation of aluminium tolerance genes. *Trends in Plant Science* **17**: 341–348.

Famoso A, Clark R, Shaff J, Craft E, McCouch SR, Kochian L. **2010**. Development of a novel aluminum tolerance phenotyping platform used for comparisons of cereal Al tolerance and investigations into rice Al tolerance mechanisms. *Plant Physiol.* **153**: 1678–1691.

Famoso AN, Zhao K, Clark RT, Tung C-WW, Wright MH, Bustamante C, Kochian L V, McCouch SR. **2011**. Genetic architecture of aluminum tolerance in rice (*Oryza sativa*) determined through genome-wide association analysis and QTL mapping. *PLoS genetics* **7**: e1002221.

Foy CD. **1988**. Plant adaptation to acid , aluminum - toxic soils. *Communications in Soil Science and Plant Analysis* **19**: 959–987.

Furukawa J, Yamaji N, Wang H, Mitani N, Murata Y, Sato K, Katsuhara M, Takeda K, Ma JF. **2007**. An aluminum-activated citrate transporter in barley. *Plant & cell physiology* **48**: 1081–91.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, *et al.* 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology* **29**: 644–52.

Grisel N, Zoller S, Kunzli-Gontarczyk M, Lampart T, Munsterkotter M, Brunner I, Bovet L, Metraux J-P, Sperisen C. **2010**. Transcriptome responses to aluminum stress in roots of aspen (*Populus tremula*). *BMC Plant Biology* **10**: 185.

Haling RE, Simpson RJ, Culvenor RA, Lambers H, Richardson AE. **2010**. Effect of soil acidity, soil strength and macropores on root growth and morphology of perennial grass species differing in acid-soil resistance. **34**: 444–456.

Hedberg I. **1990**. Morphological, cytotaxonomic and evolutionary studies *in Anthoxanthum odoratum* L. S. Lat. - A critical review. *Sommerfeltia*: 97–107.

Hiradate S, Nomoto K, Iwashita T, Matsumoto H, Danchi TN. **1997**. Interna1 detoxification mechanism of Al in hydrangea. *Plant Pysiol.* **113**: 1033-1039.

Horst WJ, Wang Y, Eticha D. **2010**. The role of the root apoplast in aluminium-induced inhibition of root elongation and in aluminium resistance of plants: a review. *Annals of botany* **106**: 185–97.

Houde M, Diallo AO. **2008**. Identification of genes and pathways associated with aluminum stress and tolerance using transcriptome profiling of wheat near-isogenic lines. *BMC Genomics* **9**: 400.

Huang C-F, Yamaji N, Chen Z, Ma JF. **2011**. A tonoplast-localized half-size ABC transporter is required for internal detoxification of aluminum in rice. *The Plant Journal* **69**: no–no.

Huang CF, Yamaji N, Mitani N, Yano M, Nagamura Y, Ma JF. **2009**. A bacterial-type ABC transporter is involved in aluminum tolerance in rice. *The Plant Cell*. **21**:655-667

Jones DL. **1998**. Organic acids in the rhizosphere – a critical review. *Plant and Soil* **205**: 25–44.

Kochian L V., Piñeros M a., Hoekenga O a. 2005. The physiology, genetics and molecular biology of plant aluminum resistance and toxicity. *Plant and Soil* **274**: 175–195.

Kumari M, Taylor G, Deyholos M. **2008**. Transcriptomic responses to aluminum stress in roots of *Arabidopsis thaliana*. *Molecular Genetics and Genomics* **279**: 339–357.

Liszkay A, Zalm E Van Der, Schopfer P. **2004**. Production of reactive oxygen intermediates ($O^{2\cdot-}$, H_2O_2 , and $^{\cdot}OH$) by maize roots and their role in wall loosening and elongation growth. Journal name is missing **136**: 3114–3123.

Livak KJ, Schmittgen TD. **2001**. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* **5**: 402–8.

Ma JF. **2000**. Role of Organic Acids in Detoxification of Aluminum in Higher Plants. *Plant Cell Physiol.* **41**: 383–390.

Ma JF. **2007**. Syndrome of aluminum toxicity and diversity of aluminum resistance in higher plants. *International review of cytology* **264**: 225–52.

Magalhaes JV. **2006**. Aluminum tolerance genes are conserved between monocots and dicots. *Proceedings of the National Academy of Sciences* **103**: 9749–9750.

Magalhaes J V, Garvin DF, Wang Y, Sorrells ME, Klein PE, Schaffert RE, Li L, Kochian L V. **2004**. Comparative mapping of a major aluminum tolerance gene in sorghum and other species in the poaceae. *Genetics* **167**: 1905–1914.

Magalhaes V J, Liu J, Guimaraes CT, Lana UGP, Alves VMC, Wang Y-H, Schaffert RE, Hoekenga OA, Pineros MA, Klein PE, *et al.* 2007. A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. *Nature Genetics* **39**: 1156.

Maron LG, Kirst M, Mao C, Milner MJ, Menossi M, Kochian L V. **2008**. Transcriptional profiling of aluminum toxicity and tolerance responses in maize roots. *New Phytologist* **179**: 116–28.

Maron LG, Piñeros M a, Guimarães CT, Magalhaes JV, Pleiman JK, Mao C, Shaff J, Belicuas SNJ, Kochian L 2010. Two functionally distinct members of the MATE (multi-drug and toxic compound extrusion) family of transporters potentially underlie two major aluminum tolerance QTLs in maize. *The Plant Journal* **61**: 728–40.

McNeilly T, Antonovics J. **1968**. Evolution in closely adjacent plant populations IV. Barriers to gene flow. *Heredity* **23**: 205–218.

Metali F, Salim KA, Burslem DFRP. **2011**. Evidence of foliar aluminium accumulation in local, regional and global datasets of wild plants. *New Phytologist*: **3**: 637-49.

Mittler R. **2002**. Oxidative stress, antioxidants and stress tolerance. *Trends Plant Sci* **7**: 405–410.

Shen R, Ma JF, 2001. Distribution and mobility of aluminium in an Al-accumulating plant, *Fagopyrum esculentum* Moench. *Journal of experimental botany.* **52**: 1683–1687.

Nishitani K, Tominaga R. **1992**. Endo-xyloglucan transferase, a novel class of glycosyltransferase that catalyzes transfer of a segment of xyloglucan molecule to another xyloglucan molecule. *The Journal of biological chemistry* **267**: 21058–64.

Pineros MA, Geraldo MAC, Lyza GM, Sangbom ML, Marcelo M, Leon VK. **2008**. Not all ALMT1-type transporters mediate aluminum-activated organic acid responses: the case of *ZmALMT1,* an anion-selective transporter. *The Plant Journal* **53**: 352–367.

Piñeros MA, Shaff JE, Manslank HS, Alves VMC, Kochian L V. **2005**. Aluminum resistance in maize cannot be solely explained by root organic acid exudation. A comparative physiological study. *Plant Physiology* **137**: 231–241.

Pitcher D, Russo MJ, Conservancy TN. **1988**. Elemant stewardship abstract for *Anthoxanthum odoratum,* Sweet Vernal Grass. Arlington: The Nature Conservancy.

Raman H, Zhang K, Cakir M, Appels R, Garvin DF, Maron LG, Kochian L V, Moroni JS, Raman R, Imtiaz M, *et al.* 2005. Molecular characterization and mapping of ALMT1, the aluminium-tolerance gene of bread wheat (*Triticum aestivum* L.). *Genome* **48**: 781–791.

Rengel Z. **2003**. *Handbook of Soil Acidity*. New York: Marcel Dekker Inc.

Ryan PR, Delhaize E. **2010**. The convergent evolution of aluminium resistance in plants exploits a convenient currency. *Functional Plant Biology* **37**: 275.

Sasaki T, Matsumoto H, Yamamoto Y, Ezaki B, Katsuhara M, Ahn SJ, Ryan PR, Delhaize E. **2004**. A wheat gene encoding an aluminum-activated malate transporter. *The Plant Journal* **37**: 645–653.

Sawaki Y, Iuchi S, Kobayashi YY, Ikka T, Sakurai N, Fujita M, Shinozaki K, Shibata D, Kobayashi M, Koyama H. **2009**. STOP1 regulates multiple genes that protect arabidopsis from proton and aluminum toxicities. *Plant physiology* **150**: 281–94.

Schluter D, Clifford E a, Nemethy M, McKinnon JS. **2004**. Parallel evolution and inheritance of quantitative traits. *The American naturalist* **163**: 809–22.

Schopfer P. **2001**. Hydroxyl radical-induced cell-wall loosening in vitro and in vivo: implications for the control of elongation growth. *The Plant Journal* **28**: 679–88.

Silvertown J, Poulton P, Johnston E, Edwards G, Heard M, Biss PM. **2006**. The Park Grass Experiment: its contribution to ecology. *Journal of Ecology* **94**: 801–814.

Snaydon RW. **1970**. Rapid population differentiation in a mosaic environment. I. The response of *Anthoxanthum odoratum* populations to soils. *Evolution* **24**: 257–269.

Snaydon RW, Davies MS. **1972**. Rapid population differentiation in a mosaic environment. II. Morphological variation in *Anthoxanthum odoratum*. *Evolution* **26**: 390–405.

Snaydon RW, Davies MS. **1976**. Rapid population differentiation in a mosaic environment IV. Populations of *Anthoxanthum odoratum* at sharp boundaries. *Heredity* **37**: 9–25.

Taylor P, Osaki M, Watanabe T, Tadano T. **1997**. Beneficial effect of aluminum on growth of plants adapted to low pH soils. *Soil Science and Plant Nutrition* **43**: 551–563.

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. **2012**. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols* **7**: 562–78.

Wenzl P, Patiño GM, Chaves a L, Mayer JE, Rao IM. **2001**. The high level of aluminum resistance in signalgrass is not associated with known mechanisms of external aluminum detoxification in root apices. *Plant physiology* **125**: 1473–84.

Xia J, Yamaji N, Kasai T, Ma JF. **2010**. Plasma membrane-localized transporter for aluminum in rice. *Proceedings of the National Academy of Sciences* **107**: 18381–18385.

Yamaji N, Huang CF, Nagao S, Yano M, Sato Y, Nagamura Y, Ma JF. **2009**. A Zinc Finger Transcription Factor ART1 Regulates Multiple Genes Implicated in Aluminum Tolerance in Rice. *The Plant Cell Online* **21**: 3339–3349.

Yang JL, Li YY, Zhang YJ, Zhang SS, Wu YR, Wu P, Zheng SJ. **2008**. Cell wall polysaccharides are specifically involved in the exclusion of aluminum from the rice root apex. *Plant physiology* **146**: 602–11.

Yokosho K, Yamaji N, Ma JF. **2011**. An Al-inducible MATE gene is involved in external detoxification of Al in rice. *The Plant journal : for cell and molecular biology* **68**: 1061–9.

Zheng Y, Zhao L, Gao J, Fei Z. **2011**. iAssembler: a package for de novo assembly of Roche-454/Sanger transcriptome sequences. *BMC bioinformatics* **12**: 453.

Zhu XF, Shi YZ, Lei GJ, Fry SC, Zhang BC, Zhou YH, Braam J, Jiang T, Xu XY, Mao CZ, *et al.* 2012. XTH31, encoding an in Vitro XEH/XET-active enzyme, regulates aluminum sensitivity by modulating *in vivo* XET action, cell wall xyloglucan content, and aluminum binding capacity in *Arabidopsis*. *The Plant cell* **24**: 1–18.

CHAPTER 4

Candidate gene analysis indicates a complex genetic basis of rapid evolution of soil aluminum tolerance in sweet vernal grass at the Park Grass Experiment

Abstract

Rapid evolution is known to enable organisms to persist in the face of environmental change. The question remains whether rapid evolution occurs by genetic mechanisms that are similar across species or by divergent pathways. In this study we examined the genetic basis of rapid evolution of tolerance to soil aluminum in *Anthoxanthum odoratum* in response to soil manipulations over the past century at the Park Grass Experiment. We genotyped markers in candidate Al tolerance loci across plants from different soil treatments in order to search for association between genotype at these loci and the Al tolerance phenotype. We find some support for the role of one locus known from crop grasses, STAR1, in the evolution of Al tolerance at PGE. However the quantitative distribution of this phenotype across the site also supports a potential role of one or more novel pathways in the evolution of tolerance that prompt further study.

Introduction

Select cases of rapid evolution have been identified in natural settings and provide some of the best evidence of evolution in action (McNeilly & Antonovics, 1968; Cook,

1976; Snaydon & Davies, 1982; Smith & Bernatchez, 2008; Cook & Saccheri, 2013). In modern ecological genetics we continue to parse out the molecular mechanisms underlying evolution in these systems (e.g. van't Hof et al. 2011). Modern sequencing technology allows us the power to address long-standing questions in evolutionary biology such as whether evolution occurs by the same or different genetic pathways in different environments and in different species, and what types of genetic variation are directly acted on by natural selection. In this study we re-examined an historic, case of rapid phenotypic evolution at the Park Grass Experiment and explored the genetic basis of that evolution using a candidate gene approach.

The Park Grass Experiment (PGE) in Harpenden, UK, is the longest-running ecological experiment in the world (Silvertown *et al.,* 2006). Originally begun as a study of pasture productivity and management, 2.8 hectares of land in a natural haymeadow have been systematically subdivided into plots and continually treated with controlled fertilizer treatments since 1856 and lime treatments since 1903. Treatment with ammonium sulfate fertilizers has caused gradual acidification of the soil while liming has countered acidification and adjusted pH upward. The result of fertilizer application and subsequent soil acidification, as in many modern agronomic settings, has been an increase in the availability of phytotoxic Al^{3+} ions in the low pH soil subplots. One early study demonstrated that a common grass at the site, *Anthoxanthum odoratum* (sweet vernal grass) has evolved locally adaptive variation in tolerance to the highly variable levels of soil aluminum across the site (Davies & Snaydon, 1973; Snaydon & Davies, 1976). Flowering time differences in *A. odoratum* between plots of different Al content have also been observed (Davies & Snaydon, 1976; Silvertown *et al.*, 2005), but a recent

study demonstrated there is still extensive gene flow between subplots (Freeland *et al.*, 2012). The population of *A. odoratum* at PGE is phenotypically differentiated with regard to soil Al tolerance, but lacks population structure that might confound the search for adaptive genetic variation in the trait. This presents an opportunity to explore the genetic basis of rapid evolution at the site.

Al tolerance is an important agronomic as well as ecological trait and has been studied at the genetic level only in cultivated species. Al primarily causes damage at the root-tip, binding with components of the cell wall, causing lesions and severe inhibition of growth as roots expand through the soil (Kochian *et al.*, 2005; Ma, 2007). In grasses such as sorghum (Magalhaes V *et al.*, 2007) and the Triticeae (wheat, barley, and rye (Magalhaes *et al.*, 2004; Sasaki *et al.*, 2004; Fujii *et al.*, 2012)) Al tolerance has near-Mendelian inheritance patterns and is linked to increased efflux of organic acids from the root-tip. These acids in turn bind toxic Al^{3+} ions in the root apoplast and soil, inhibiting their absorption (Ryan & Delhaize, 2001). Transposon insertions in the promoter region of key organic acid transport genes in at least two species have been shown to strongly influence the level of tolerance (Raman *et al.*, 2005; Magalhaes V *et al.*, 2007). In rice and corn, organic acid transport genes also contribute to Al tolerance but explain less of the overall variation in the trait. In these species, Al tolerance has more complex polygenic inheritance (Krill *et al.*, 2010; Famoso *et al.*, 2011). In wild grasses Al tolerance has been measured in several species (Foy, 1988; Wheeler, 1995), but the genetic basis of the trait has not been explored.

Anthoxanthum odoratum is a temperate out-crossing grass that thrives in acidic, high Al soils and has the ability to produce root growth at Al^{3+} ion activities far beyond

the cultivated grasses and many other wild herbaceous plants (Chapter 2). In this study, we tested the Al tolerance of both adult plants and seedlings from PGE in order to characterize phenotypic differentiation across soil types. We then tested for associations between phenotype and genotype at a suite of candidate loci that were identified based on expression patterns across genotypes in response to Al (Chapter 3). Associations can be detected if single nucleotide polymorphisms (SNPs) in these transcripts are genetically linked with nearby polymorphisms that effect gene regulation. In this case we predict that SNPs in transcripts linked to variation in Al tolerance will be differentiated across high and low Al soil types at PGE, and that tolerance of individual plants would be associated with the presence/absence of those markers.

Methods

The Park Grass Experiment. PGE is a large hayfield that has been subdivided into experimental treatment subplots in a factorial design (Silvertown et al. 2006, Fig. 1). Annual soil applications of macronutrients such as sodium nitrate and ammonium sulfate fertilizers are crossed with four liming treatments, which began in 1903. Liming is applied at four levels (A through D, Fig. 1), enough to amend the soil to pH 7, 6, and 5, respectively, with the last subplot (D) left un-limed. The unlimed subplots at PGE have become acidified to near pH 4 primarily as a result of the fertilizer application (Van Bergen *et al.*, 1998). Plot 3 at the experiment is a control plot where liming treatments have been applied but no nutrients have been added.

Soils are regularly tested and archived at PGE. Archived soils have been sampled to a depth of 23 cm, air dried, milled to 2 mm, and stored at room temperature. We sub-sampled the most recently available archived soils (from 2008) from the subplots where we collected seeds and tested them for nutrient and Al content (Table S1). Soil samples were tested for nutrient contents using a 0.1 M $CaCl_2$ extract, which better approximates Al availability in soils than traditional Morgan extraction (Bertsch & Bloom, 1996). Nutrients were measured in the extracts using inductively coupled plasma mass spectrometry at the Cornell Nutrient Analysis Laboratory. The soil pH varies in acidity from 7.2 to 3.7 across the plots we sampled; Al content co-varies with pH, from 0 to 836 uM (Table 1).

Plant material. *Anthoxanthum odoratum* grows on the majority of subplots at PGE, generally at lower density on more neutral soils. We collected inflorescences in July 2010 across the plots indicated in Fig. 1. We collected seeds from subplots A and D, the most neutral and the most acidic subplots respectively, from within nutrient treatments 1, 4/2, 9/2, and 3. (*A. odoratum* has low abundance on plot 9/2A, so seeds were sampled from 9B instead.) Whole inflorescences were collected from each of 16 plants within a subplot, plants were spaced at 4-5 paces apart along parallel transects and were not sampled near the edges of the subplots. Seed families (consisting of full and half sibs) were separated from each inflorescence and planted in coarse vermiculite (Whittmore Company, Inc., Lawrence, MA) followed by 3 days vernalization (Chapter 2). At 5 weeks old, one seedling per family was transplanted to a 1:1 mixture of perlite and vermiculite

in the greenhouse. For 4 families with low germination rates, plants were added from a second round of germinated seeds. This yielded an average of 13.8 plants per subplot (Table 1); a total of 108 plants were generated. Plants grew for 10 months in the greenhouse before they were tested for Al tolerance.

Because morphologically similar diploid sister species and hybrids of *Anthoxanthum sp.* occur at some locations in the native range of *A. odoratum* (Pereira *et al.*, 2007; Pimentel *et al.*, 2007), we confirmed the ploidy level of a subsample of plants at PGE. We measured the genome size of 26 adult plants (3-4 from each subplot) by flow cytometry following (Costich *et al.*, 1991; Dolezel *et al.*, 2007). All plants had close to the predicted tetraploid genome size (Grime & Mowforth, 1982; Pereira *et al.*, 2007), ranging from 12.4 to 14.1 pg.

Al tolerance testing – tillers: Al tolerance was assayed measurement of relative root growth (RRG) of vegetative tillers under Al exposure. Vegetative tillers of similar size were removed from plants in the greenhouse, the remaining roots removed, and then placed in 54 L tubs containing a modified 1/10 strength Long-Ashton solution at pH 4.0 (Chapter 2). Each tub contained 45 tillers. We submitted hydroponic solution sampled throughout the experiment to verify that no micro- or macronutrient was exhausted during the growing period (data not shown). Tillers were pruned to 3-4 fully expanded leaves and long leaves were trimmed to 15 cm (to equalize leaf area between tillers). Tillers were tested in a split-plot design, with the same compliment and positioning of genotypes replicated in pairs of tubs (3 pairs per experimental trial). One tub in each pair was subjected to a control treatment and the other an Al exposure treatment. After 5 days

acclimation, the length of the longest root (L1) and weight of each tiller was measured, and the hydroponic solution was replaced with fresh solution. Length of the longest root was measured again after 4 days growth (L2), and the solution replaced with either fresh solution (control treatment) or solution containing 950 uM Al (equivalent to an Al^{3+} ion activity (availability) level of 300 uM (Chapter 2)**.** The length of the longest root was measured again after 4 days (L3). We calculated Al tolerance of each tiller as relative root growth (RRG): the difference between total growth in the first 4-day period (P1 = L2-L1) and the second period (P2 = L3-L2). In the Al treatments lower values of RRG indicate lower Al tolerance. The entire experiment was repeated 7 times (trials) to generate 4-8 replicate measurements (average 5.5) per adult plant genotype in both the control and Al treatments (control n=580 tillers total, Al treatment n=600 tillers total). Variation between temporal blocks (trials) and spatial blocks (tub pairs) was accounted for in linear models predicting average genotypic Al tolerance.

Al tolerance testing – seedlings: Seeds from 5-11 families per PGE subplot (1-3 seedlings per family, per treatment) were used to test Al tolerance. Groups of four seeds per family were soaked overnight, sterilized in 75% ethanol for 10 minutes, and germinated on solidified nutrient solution (Chapter 2). After 3 weeks, seedlings were transplanted to tubs containing 9 L of hydroponic solution at pH 4.0, 20 seedlings per tub, arranged in a split plot design. Sampling of plant genotypes was balanced by PGE subplot as described above for tillers. The solutions were replaced after 10 days acclimation and the length of the longest root, height, and the number of leaves of each seedling was measured. Seedlings were then treated and measured as described above for

tillers. All seedlings were tested in a single experiment. We excluded data from 33 seedlings that did not grow or died at the start of the experiment due to transplant stress (18 were from the Al treatment, 14 were from the control treatment). We used measurements from a total of 4-13 seedlings per PGE subplot per treatment (n=60 total control seedlings, n=64 Al treated seedlings).

Statistical Analysis: Average Al tolerance for adult plants from each soil subplot at PGE was calculated as the least-squared mean from a linear model with factors: plot (1, 4/2, or 9/2), soil-type (A or D subplot), plant genotype (greenhouse individual) nested within plot, baseline growth rate (P1, above), treatment (control or +Al), treatment by soil-type interaction, temporal block, and spatial block (tub pair) nested within temporal block. All factors were fixed except for plant genotype. Least-squared means for subplots 3A and 3D were calculated separately with the same model minus the plot effect (Fig. 2A and 2B). The tolerance of seedlings from high and low Al soils at PGE was calculated as the least-squared means from a linear model with factors: plot, soil type, baseline growth rate (P1), treatment, spatial block (tub) nested within treatment, and a treatment by soil type interaction (Fig. 2C and 2D). All factors were fixed.

SNP genotyping: SNPs were identified within expressed sequences from Al-treated root tip cDNA libraries of one Al tolerant and one Al sensitive plant (Chapter 3). The sequenced individuals came from subplots 4A and 4D, respectively. SNPs in differentially regulated candidate transcripts were identified using read alignments to the transcriptome in Integrated Genomics Viewer (Thorvaldsdóttir *et al.*, 2012). We also

identified SNPs in a set of randomly chosen transcripts (termed 'population genetic transcripts' from here on). We designed primers for genotyping these markers on the Sequenom MassArray platform at the Cornell Core Laboratories Center using Assay Designer 4.0.

We extracted genomic DNA from young leaf tissue in 96-well plates using cTAB phenol-chloroform extraction. Tetraploid SNP genotypes were called based on peak signal to noise ratios using Typer 4.0 and custom R scripts designed by M. Sylvester at Sequenom Inc (unpublished data). Genotyping was based on a clustering algorithm designed by (Fujisawa *et al.*, 2004). Data from SNP markers with low amplification or no diversity, and plant DNAs with a low frequency of successful SNP calls (n=3) were discarded. The final data set included data for 50 SNPs from within 31 candidate transcripts, and 94 SNPs from 94 population genetic transcripts across 105 individuals.

Genotype-phenotype association analysis. Population genetic statistics for each marker and subplot were calculated using the program AUTOTET (Thrall & Young, 2000). F_{st} between soil types (high and low Al soil plots) was calculated based on the expected heterozygosity among all plants from the low Al soils (including subplots 3A and 3D), expected heterozygosity among all plants from the high Al soils, and the expected heterozygosity across all plants together. For some markers, measured F_{st} across soil types was slightly negative due to stochasticity of random sampling within subpopulations. To examine population structure, we analyzed data from the 94 population genetic SNPs using the program STRUCTURE (Pritchard *et al.*, 2000). We used a burn-in period of 10,000 iterations and ran the analysis a further 10,000 iterations

at k=2 and k=4 groups. Because no population structure was detected in either analysis, we conducted further analysis of association between genotypes and phenotype without population coefficients. We tested for association between each marker and Al tolerance using Fisher's exact tests (FETs). Genotypes with the 20 highest least-squared mean Al tolerance were designated the 'tolerant' phenotype and those with the 20 lowest were designated the 'sensitive' phenotype. Exact tests and examined association between this categorical classification and genotype at each SNP marker. For candidate SNP markers with p-values in the top 5% of the distribution we tested the significance between genotypes using phenotypic data across all genotypes with a Kruskal-Wallis non-parametric test.

Results

Al tolerance of adult plants and seedlings. Treatment with Al had a significant effect on average RRG of plants from both high and low Al soils at PGE. Interestingly, root growth of seedlings was on average more resistant to Al exposure than that of tillers. Both average tiller and seedling RRG measured in control treatments was near zero (root growth was uninhibited) but in Al treatments root growth was reduced on average by about 2 cm over 4 days (Fig. 2). For tillers, variation in RRG between clonal replicates of the same genotype was similar in magnitude and distribution in the control and Al treatments (Fig. 2). The variance in tiller RRG measurements for each genotype was also positively correlated between control and Al treatments (p=0.003), indicating that variability in RRG is likely to be biologically based rather than a result of environmental or experimental variation.

The relationship between baseline growth rate (P1), RRG in control treatments, and Al tolerance was consistent at the seedling and adult life stages. In tillers, weight and baseline growth rate (P1) were significant predictors of RRG in both control and Al treatments. In seedlings baseline growth rate was positively correlated with number of leaves (corr=0.37) and height (corr=0.40) and was also a significant predictor of RRG. This is because at both life stages there was a negative correlation between baseline growth rate and RRG, and this relationship was enhanced in the presence of Al (Fig. 3A vs 3B, 3C vs. 3D). Both faster growing tillers and larger faster growing seedlings tended to have more sensitivity to Al exposure.

Overall, both adult plants and seedlings from high Al acid soils at PGE had higher Al tolerance than those from low Al neutral soils (treatment by soil type interaction, p=0.02 and 0.10 respectively, Fig. 4A and 4C). While differences in tolerance are small, further support for the reality of this pattern is evident in that RRG did not differ significantly between soil types in the absence of Al exposure (control treatments) and also did not differ between plants from the two control soil subplots with or without exposure to Al (Fig. 4B and 4D). Also, within each pair of high/low Al subplots within soil nutrient treatments at PGE, tillers and seedlings from the high Al subplot in each pair consistently have higher Al tolerance (Fig. 5).

Population genetic structure. Data from the set of 94 population genetic SNPs, did not show evidence of any population structure across nutrient treatment plots at PGE or between high and low Al soils (Fig. 6). Average F_{st} across high and low Al soils was also low at 0.03 (Fig. 8). These results match the expectation of high gene flow between

plants from different soils because of their close proximity and the mode of out-crossing wind pollination in *A. odoratum*.

Genotype-phenotype association. We also genotyped 50 SNPs located in the coding regions of 31 candidate Al tolerance transcripts (Table 2). The level of differentiation (F_{st}) between plants on different soil types for the majority of candidate SNPs was within the range of differentiation for the population genetic SNPs (Fig. 8). However, F_{st} for five candidate SNPs fell above the upper 5th percentile of the distribution of population genetic marker F_{st}. The highest F_{st} outlier was for a marker in a transcript with high similarity to the Al-activated malate transporter ALMT1 (unigene UN11139, $F_{st} = 0.117$). The next two highest F_{st} outliers were associated with SNPs in putative homologs of STAR1 and STAR2 in rice, which are hypothesized to function in modification of the cell wall.

Our ability to detect statistically significant genotypic differences in Al tolerance was reduced by small sample size, partly a limitation of the small spatial scale of PGE itself. In addition, if some of the SNPs represent differences between two similar but parologous loci, the relationship between genotype and phenotype would be difficult to decipher. Despite this, among our association tests of SNP genotype with phenotype some SNPs showed patterns suggestive of linkage with Al tolerance. The unigene with the largest phenotypic differences between tetraploid SNP genotypes was a methionyl tRNA synthetase-like transcript (exact test, p=0.01) followed by a pra1 family disease resistance protein (exact test, p=0.03) (Fig. 9, Table 2). However, for neither SNP was there a clear pattern of increasing or decreasing tolerance with increasing homozygosity,

as would be predicted if there was an additive phenotypic effect of having multiple copies of the SNP at a single tetraploid locus. We did find a weak trend between homozygosity of SNP markers and Al tolerance in four transcripts: MGT1, STAR1 and STAR2, and a MATE-family unigene (Fig. 9, Table 2). Among these, the SNP in STAR1 has both high F_{st}, marginally significant association with phenotype, and increasing Al tolerance with increasing homozygosity. It is the strongest candidate SNP for direct linkage with genetic polymorphisms that have been influenced by natural selection at PGE.

Discussion

Across soil types at PGE variation in Al tolerance in *A. odoratum* is quantitatively distributed (Fig. 2). The lack of bimodality in phenotype suggests that the source of genetic variation in Al tolerance is polygenic in nature, contrary to what has been observed between some cultivars of sorghum, wheat, and oat. Despite this, we did detect differences on average in the Al tolerance of plants from different soil types, and these differences are locally adaptive in direction although weaker in strength than those observed at PGE 50 years ago (Davies & Snaydon, 1973). Davies & Snaydon may have detected stronger differentiation in tolerance because they exposed tillers to Al for a longer period of time. At the same time, their assays lacked internal controls (RRG measured without Al exposure) such that differentiation in Al tolerance may have been confounded with differentiation to low pH itself. Also, by testing an external control in this study (tillers from fertilizer-free subplots 3A and 3D) we have been able to demonstrate that Al tolerance differences are linked to soil acidification at PGE and not other differences between limed and unlimed soils.

Phenotypic trends we observed are consistent with what might be predicted based on the action of Al at root cell walls. Faster growing seedlings and tillers are more sensitive to Al. Root cell walls require flexibility as they are continually broken down and reformed during root expansion during growth. Thus rigidity caused by attachment of Al to pectin in cell walls is likely to cause more physical damage to faster growing roots. In seedlings we also find the reaction norms for Al tolerance are crossing (Figure 4), suggestive of a tradeoff between Al tolerance and sustained growth rate over time (non-negative RRG values) on non-acid soil. One hypothesis to explain the tradeoff is that differences in resource allocation to root growth versus Al resistance may exist between tolerant and sensitive plants. More Al tolerant plants may have higher resource allocation to fortifying cell walls, or some other constitutive mechanism of Al tolerance such as continuous organic acid release, and less to active growth of roots. (and vice versa in sensitive plants) Thus it is possible that balancing selection on the same genetic polymorphism for this mechanism across acid and non-acid soils contributes to the maintenance of genetic variation at PGE.

So what genes have facilitated evolution of the locally adaptive differences in Al tolerance at PGE? Are they similar to the genes that have been selected through breeding for acid-soil tolerance in crops? We find a marker in a transcript with high similarity to an organic acid efflux gene controlling tolerance in the Triticeae (ALMT1 in wheat, barley, and rye) was a strong outlier regarding differentiation between plants on different soil types at PGE. However the true role of this locus in facilitating local adaptation is equivocal, as we did not detect a strong relationship between tolerance and homozygosity for this marker. More promising candidates for the targets of selection at PGE are the

putative homologs of the interacting abc transport genes STAR1 and STAR2 in *A. odoratum*. In particular, one marker in STAR1 showed both high differentiation between soil types and a trend for increasing tolerance with increasing homozygosity at the locus. These two genes are vesicle-localized, involved in the transport of UDP-glucose in rice (Huang *et al.*, 2011), and may function in cell-wall modification. Selection for a cell wall related mechanism of Al exclusion would be consistent with the observed relationship between growth rate and Al tolerance. Overall, the evidence here only moderately supports the hypothesis that genetic evolution of Al tolerance in *A.odoratum* has occurred through the same pathways as in the cultivated grasses. There are over 100 transcripts of unknown function whose expression patterns indicate they are involved in Al tolerance variation at PGE (Chapter 3). Further genotyping of markers at these loci would help elucidate if any novel pathways are more strongly associated with tolerance differences than the loci from crop grasses tested here.

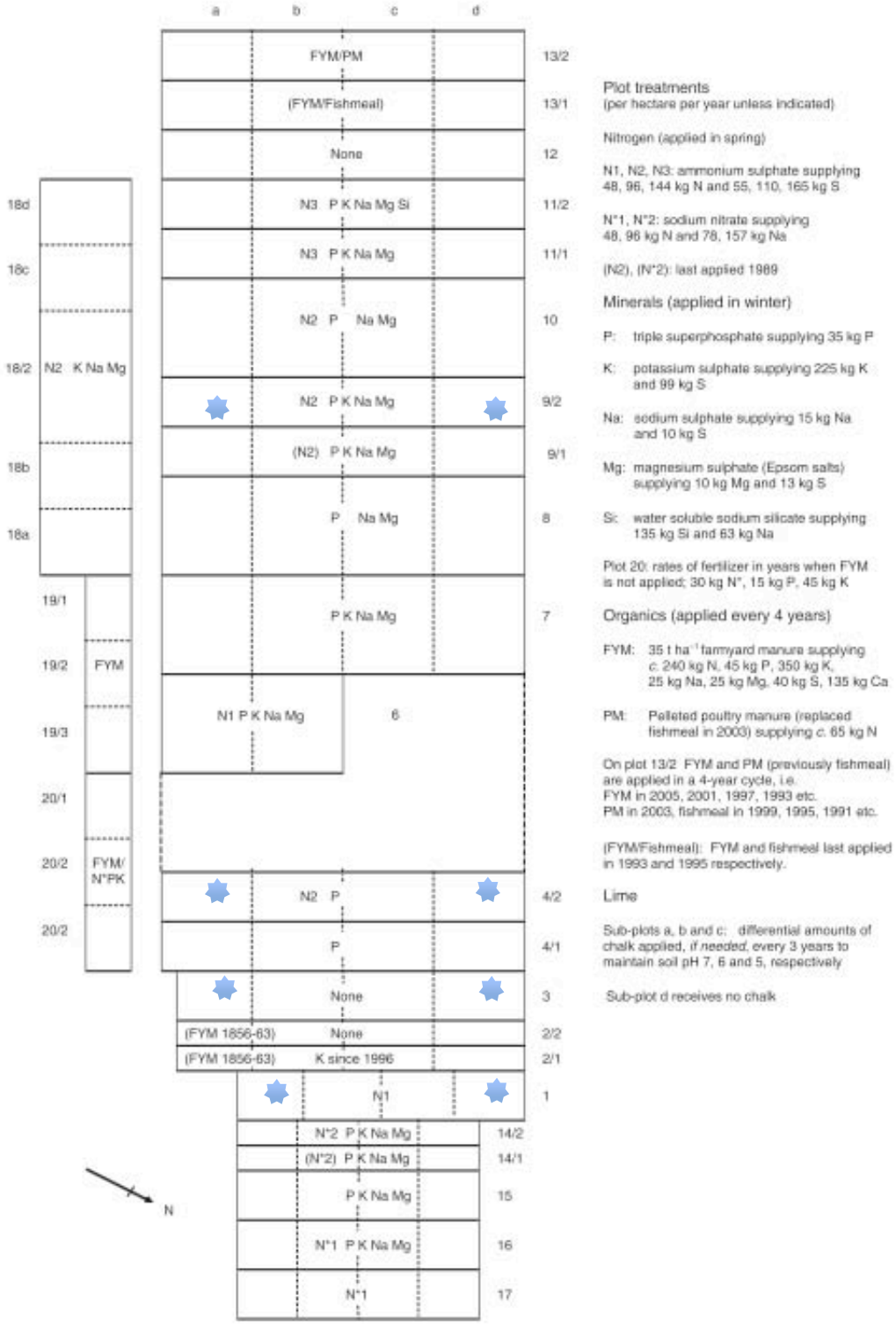
The quantitatively distributed nature of the phenotype points toward selection for multiple pathways to tolerance in *A. odoratum* and perhaps may explain the species' extreme tolerance in relation to almost all other grasses measured to date. Loci that contribute to tolerance are likely to have been independently selected from standing genetic variation over a short period of about 110 years since the start of soil pH manipulations at PGE. Outcrossing and high rates of gene flow across soil types at PGE may have helped local adaptation to evolve for a polygenic stress tolerance trait by allowing multiple adaptive alleles to effectively recombine into a single genotype that could be naturally selected on acid plots. In this way, *A. odoratum* growing at PGE may

provide a valuable future resource for the future study of the genetic basis of variation in other genetically complex traits related to soil adaptation in plants.

Acknowledgements: Special thanks to Denise Costich for help with flow cytometry; to John Storkey, Andy McDonald and Rothamsted Research for facilitating sampling at PGE. This project was funded by grants from the Cornell 3CPG and the Cornell Biogeochemistry NSF-IGERT.

FIGURES

Figure 1: Modified from Silvertown et al. 2006, Figure 2: the plot layout and current treatments of the Park Grass Experiment. ✸ indicates subplots sampled in this study.



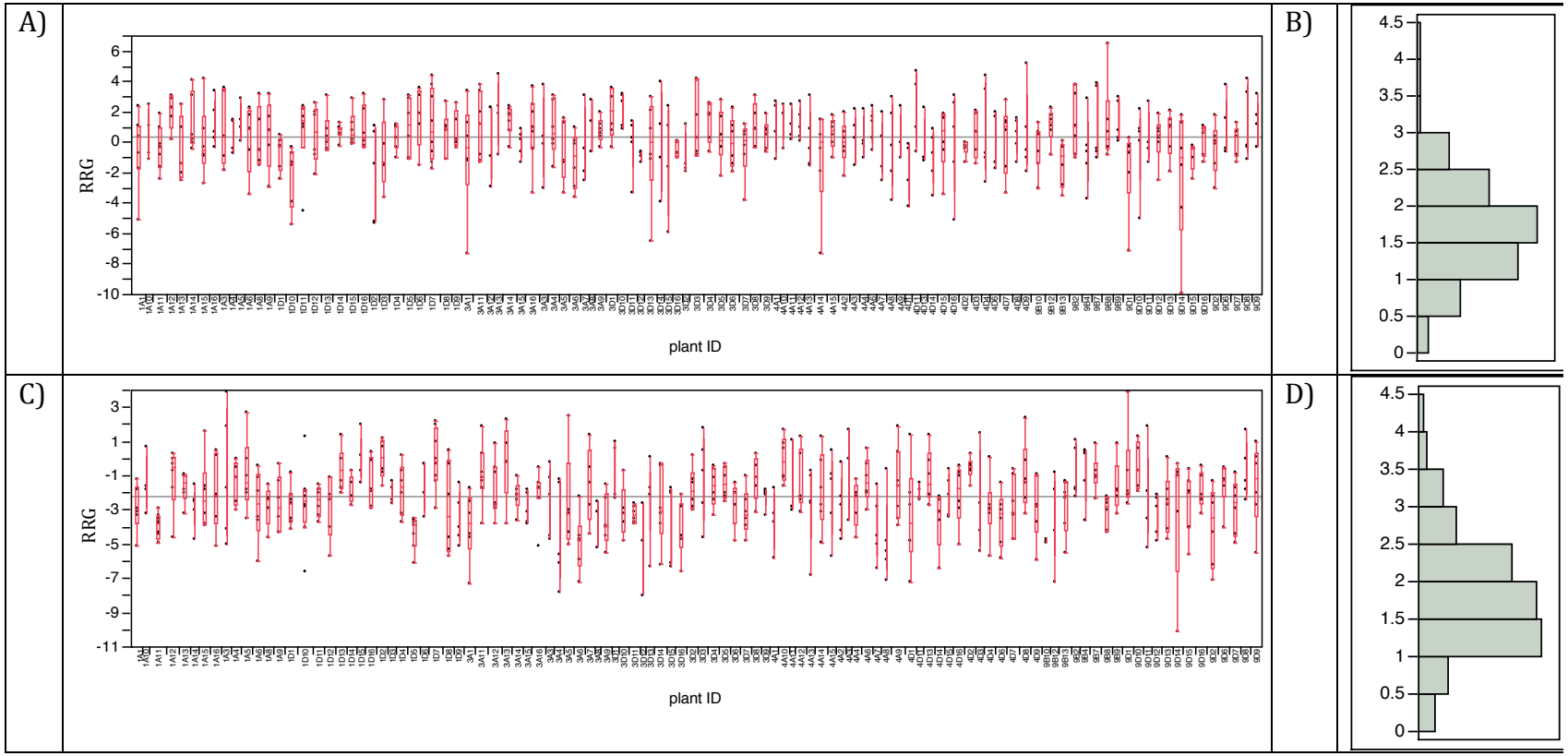
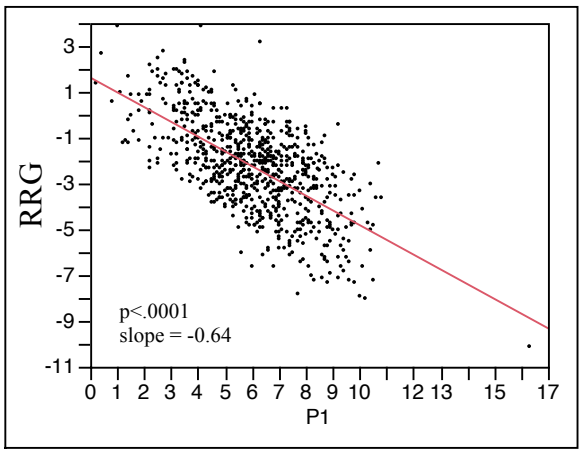


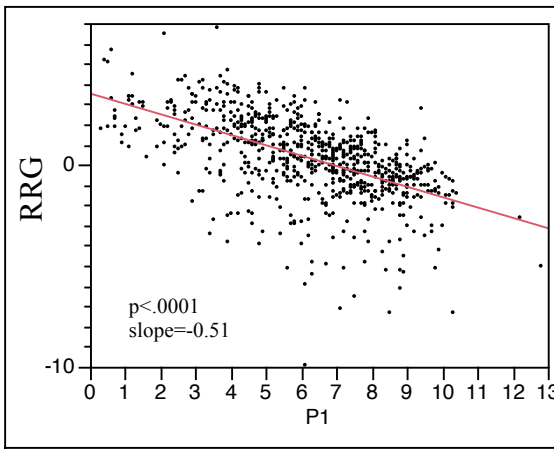
Figure 2 Variation in RRG measurements for plant genotypes. A) Control treatment RRG measurements by genotype. Boxes and lines represent quantiles. B) The distribution of standard deviations in RRG for the control treatments. C) and D) box plo and distribution of standard deviations from +Al treatments.

Figure 3. RRG versus baseline growth rate (P1). A) Al treatment, tillers B) Control treatm tillers; C) Al treatment, seedlings D) control treatment, seedlings.

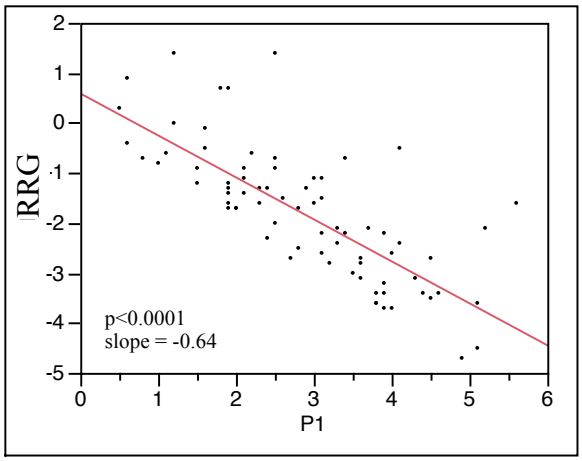
A)



B)



C)



D)

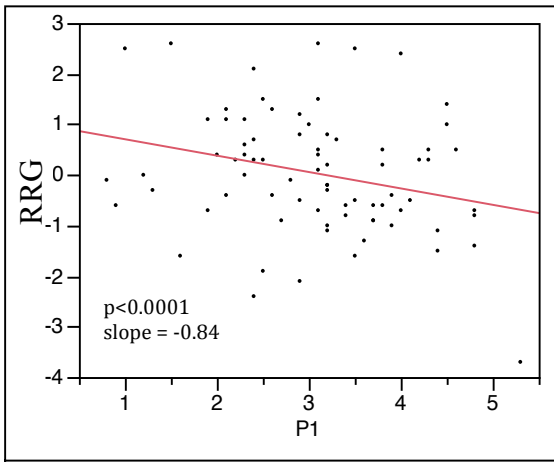
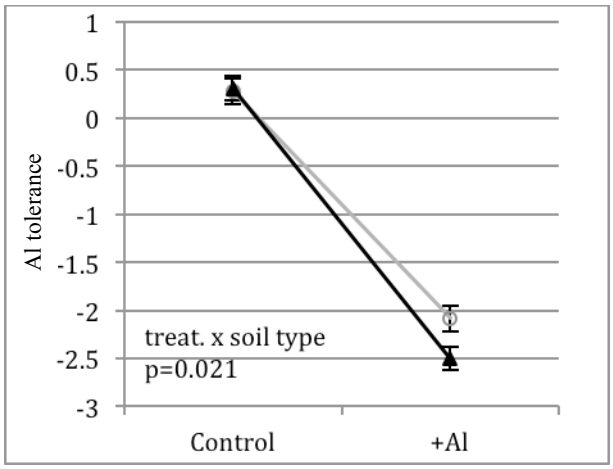
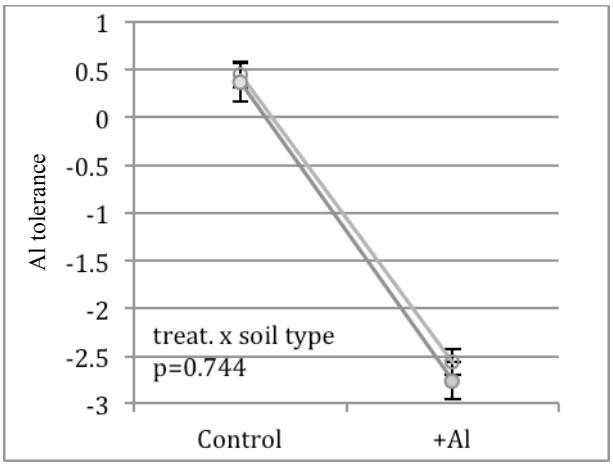


Figure 4. Reaction norms for tillers and seedlings. A) and C) dark grey triangles, plants from neutral soil; light grey circles, plants from high Al soil. B) and D) open circles, plants from subplot 3D, plants from subplot 3A. Error bars are +/- 1 S.E. Values are LSMeans from models described in the Methods.

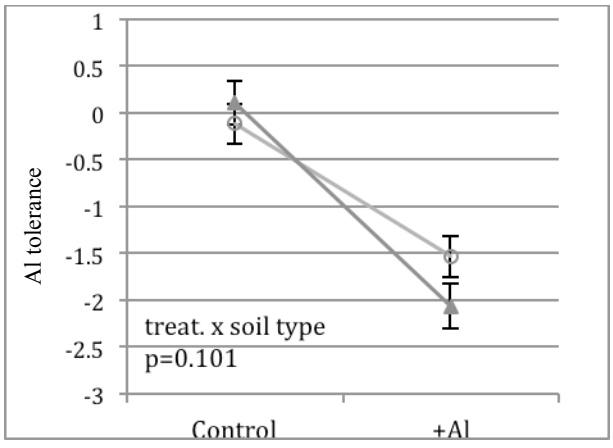
A) tillers plots (1, 4, 9)



B) tillers, control plot (3)



C) seedlings, plots (1,4,9)



D) seedlings, control plot (3)

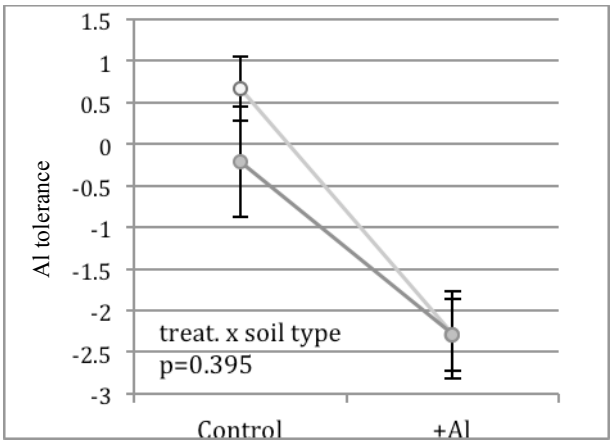
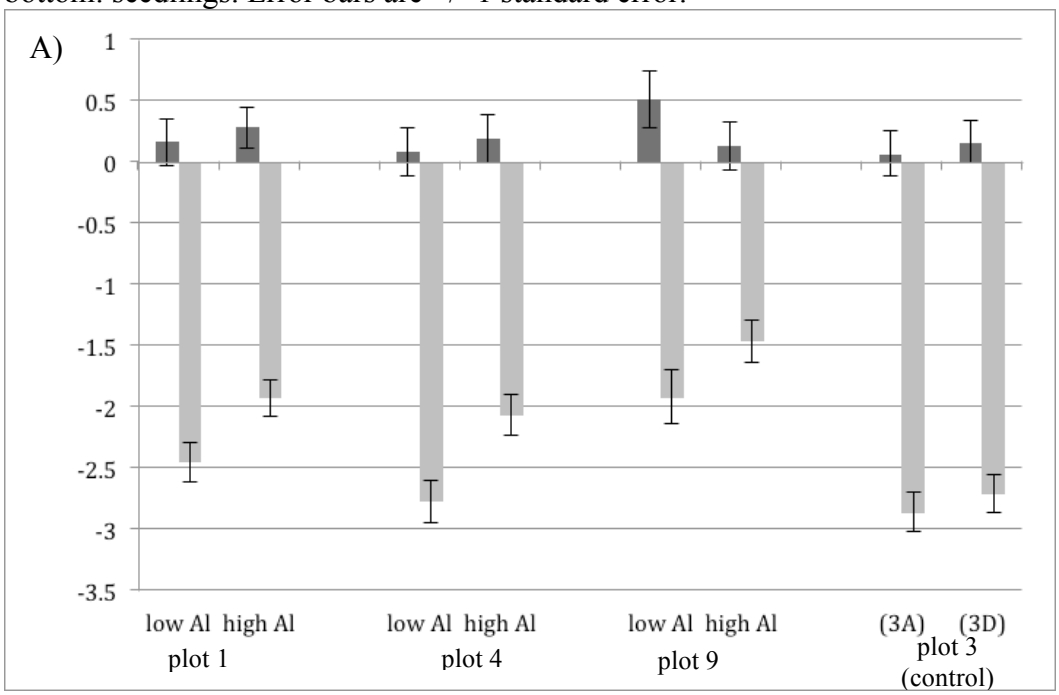


Figure 5. Al tolerance by subplot. Al tolerance is calculated as the LSMean value for each subplot from a linear model that accounts for tiller growth rate and block effects within each treatment group. Dark bars are the control treatment, light bars are the Al treatment. Top: tillers, bottom: seedlings. Error bars are +/- 1 standard error.



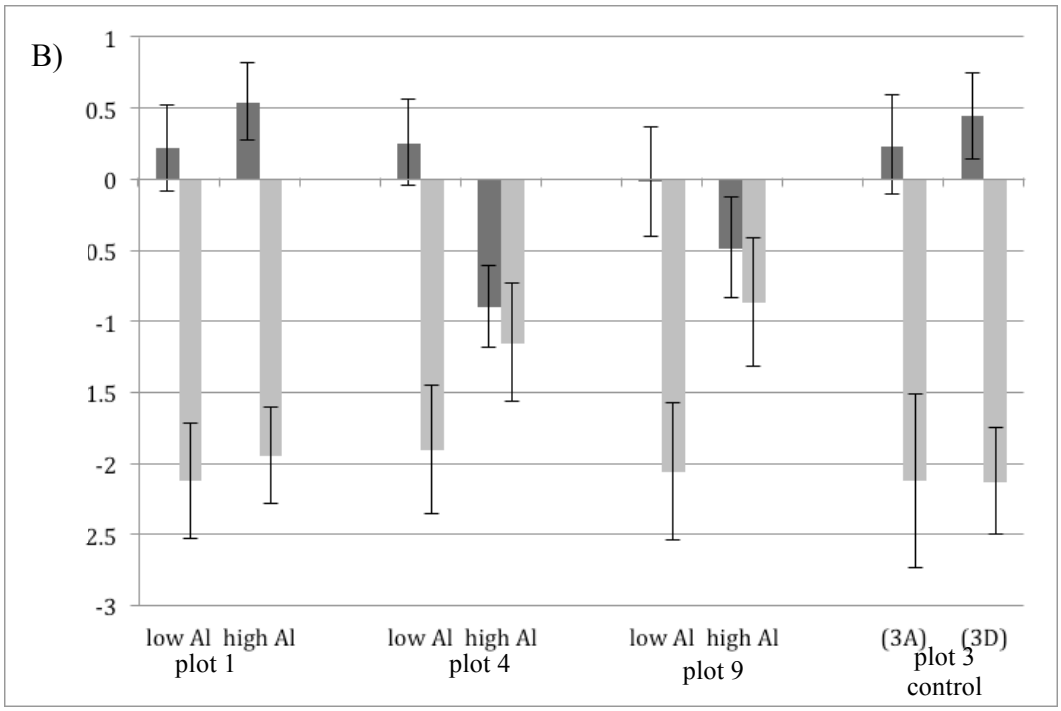


Figure. 6 Population genetic structure plots, k=2. Top: genotypes organized by soil type (0=low Al, 99=high Al). Bottom: genotypes organized by source plot.

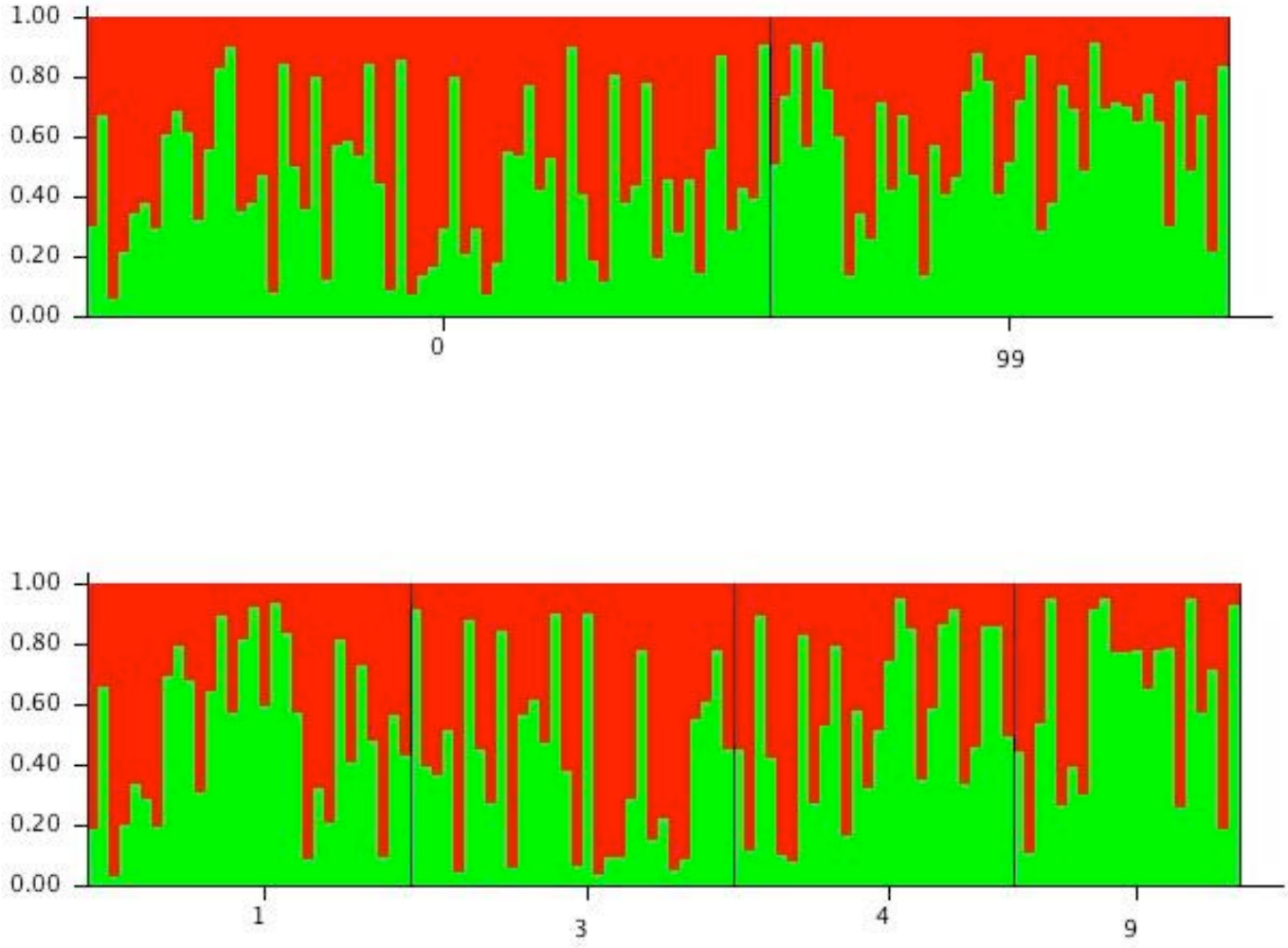


Figure 7. The distribution of pairwise F_{st} values between high and low Al soils for 94 population genetic SNPs. Dots are values for SNPs in candidate genes. Dashed line represents the 95th percentile Fst values of the distribution. ALMT1-like, UN11139; STAR1, UN11920; STAR2, UN50094.

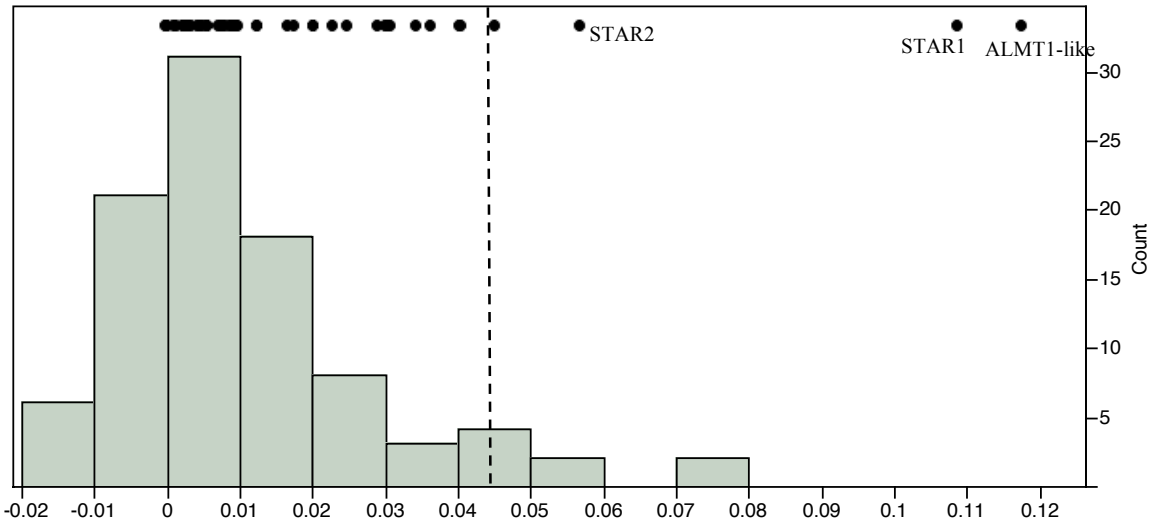


Figure 8. Frequency of significance values of association between genotypes and Al tolerance (Fisher's exact tests) across 94 population genetic markers. Dots represent values for candidate gene SNPs. Met-tRNA-like, UN13967; Pra1family protein, UN21840; STAR1, UN11920; STAR2, UN50094; MGT1, UN33791; Nrat rrp8-like protein, UN37144. Dotted line marks the 5th percentile of p-values.

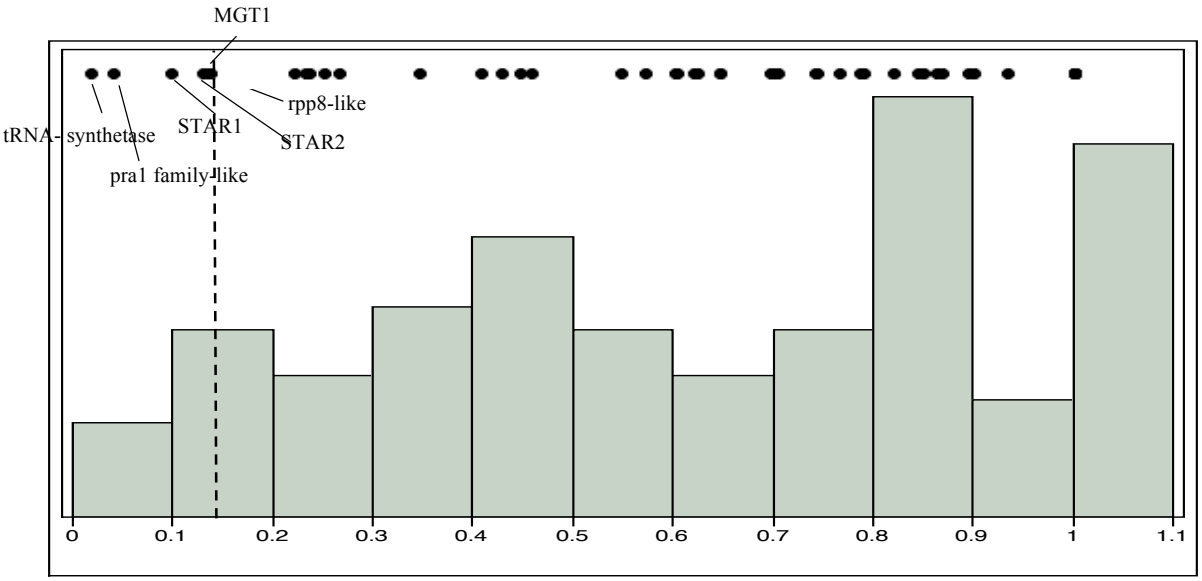


Figure 10. Genotypic differences in Al tolerance for SNPs in selected candidate transcripts. The range of phenotypic Al tolerance values is shown across plants with a different tetraploid genotype at a SNP. Shown are values for SNPs with average phenotypic values that trend in a linear direction.

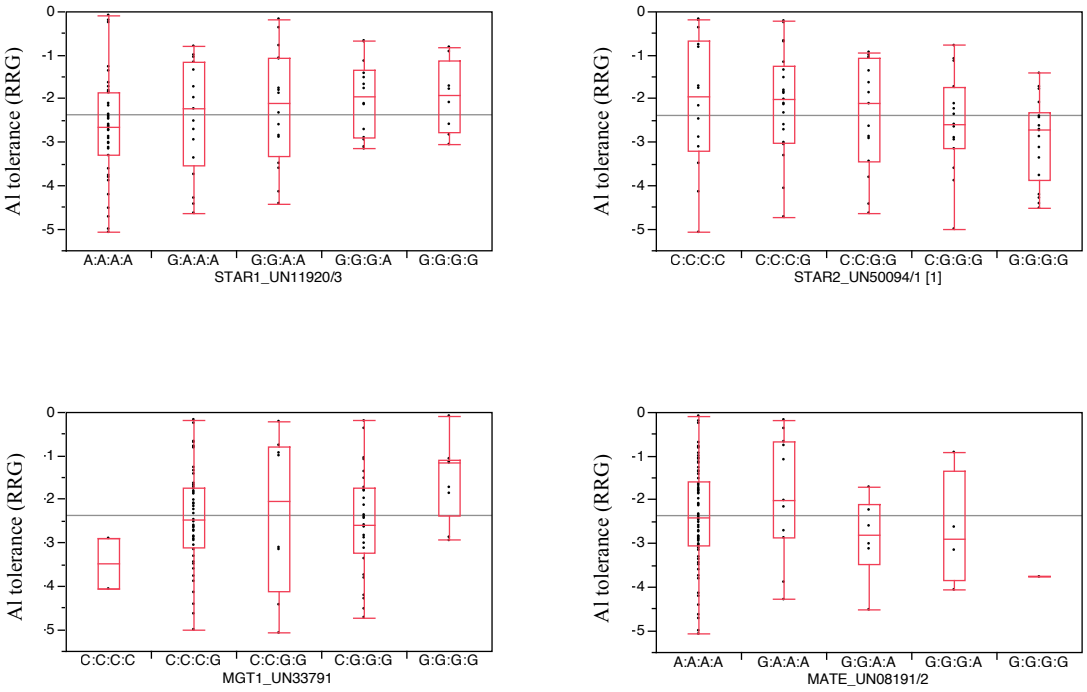


Table 1. Soil characteristics of plots sampled at PGE.

Plot	Subplot	No. plants	Soil pH	Al *uM*	Na *uM*	Mg *uM*	K *uM*	Ca *uM*	Mn *uM*	Fe *uM*	Cu *uM*	Zn *uM*
1	1A (low Al)	14	7.1	<det	468	180	125	12,276	44	5	1	7
	1D (high Al)	16	4.1	836	498	166	445	18,337	1,962	196	2	50
4/2	4A (low Al)	14	6.9	<det	428	70	276	20,282	4	5	1	1
	4D (high Al)	13	3.7	672	325	172	373	26,279	616	337	2	41
9/2	9B (low Al)	8	6.3	3	478	302	2543	28,897	35	5	1	1
	9D (high Al)	13	3.7	710	936	312	2710	2,571	210	1476	4	47
3	3A (control)	14	7.2	2	476	222	127	13,393	17	3	3	3
	3D (control)	16	5.2	6	516	397	244	39,882	1,259	3	1	14

Table 2. Candidate SNPs. Columns 2 and 3 show fold change in expression in response to Al in a sensitive and a tolerant plant. More than one SNP was genotyped for some unigenes and the statistics are listed in succession in columns 5 and 6. F_{st} values above the 95th percentile of population genetic markers, and association Fishers exact test p values below the 5th percentile are in bold.

Unigene	Sensitive fold-change	Tolerant fold-change	Description	Fst (soil type)	Association (p-val)
UN01296	2.1	6.2	metal transporter	0.002	1.00
				0.030	0.42
UN01856	1.5	0.9	similar to ATR1, serine threonine-protein kinase atr-like	0.006	0.44
UN02272	4.4	10.8	transcription factor	0.000	1.00
				-0.001	0.60
UN03316	2.0	3.3	magnesium transporter mrs2-e-like	0.000	0.62
UN08191	2.8	2.7	putative citrate efflux MATE transporter	0.003	0.25
				0.039	0.69
UN08554	4.7	2.6	Nrat1 (Nramp metal transporter)	0.000	0.78
UN09941	2.2	3.5	cytochrome p450	-0.001	0.87
				0.004	0.90
UN11139	13.6	4.9	aluminum-activated malate transporter	**0.117**	0.26
				0.009	0.89
UN11622	2.0	2.8	mate efflux family protein 5-like	-0.001	0.40
UN11920	3.2	2.9	STAR1, phosphate import atp-binding protein pstb 1	0.016	0.70
				0.0113	0.22
				0.108	**0.12**
UN13967	6.8	8.6	methionyl-trna synthetase	0.001	**0.01**
UN14079	11.2	5.4	almt1	0.011	0.64
UN15676	0.9	1.9	subtilisin-like protease-like	0.002	0.74
UN15889	15.1	22.6	eukaryotic translation initiation factor 6	0.033	0.62
				0.007	0.74

UN17334	7.3	31.1	bifunctional aminoacyl-trna	0.000	0.85
			expressed	0.001	0.62
UN21840/1/2/3	1.8	10.3	pra1 family protein d-like	0.035	0.86
				0.008	**0.03**
UN23598/9	9.0	42.7	glucosyltranserase-like	0.001	1.00
				0.029	0.70
UN26777	1.6	10.3	unknown	**0.044**	0.64
UN27341	0.3	5.2	heavy metal-associated domain-	0.001	0.93
			containing protein	0.019	1.00
UN33380	1.0	1.1	blue copper binding protein	0.006	0.70
UN33791	1.5	1.5	MGT1 (magnesium transporter)	0.000	**0.13**
UN33906	0.3	3.1	pentatricopeptide repeat-	0.016	1.00
			containing protein (Nrat-like)	0.006	0.79
UN37144	3.5	2.5	disease resistance protein rpp8-like	0.002	**0.13**
			(Nrat-like)	0.000	0.60
UN40886	10.6	365.4	ABC transporter C family member	0.022	0.23
			13-like	0.040	1.00
UN47815	5.3	2.6	aluminum-activated malate	0.028	0.23
			transporter	0.029	
UN47817	6.8	3.4	aluminum-activated malate	0.019	0.54
			transporter	0.024	
UN50096	1.6	1.0	STAR2, protein al-sensitive 3	**0.056**	0.84
				0.003	**0.09**
UN65068	0.9	0.9	ART1, c2-h2 zinc finger protein	0.002	0.58
UN71256	90.8	135.3	unknown	0.004	1.00
UN78365	2.4	1.3	ALS1-like, abc transporter	0.000	0.34
UN84836	1.5	2.0	blue copper binding protein	0.005	0.45

Literature Cited

Bertsch PM, Bloom PR. **1996**. *Methods of Soil Analysis: Part 3. Chemical Methods* (D Sparks, Ed.). Madison, WI: Soil Science Society of America and American Society of Agronomy.

Chen ZC, Yamaji N, Motoyama R, Nagamura Y, Ma JF. **2012**. Up-regulation of a magnesium transporter gene OsMGT1 is required for conferring aluminum tolerance in rice. *Plant physiology* **159**: 1624–33.

Cook LM. **1976**. The rise and fall of the carbonaria form of the peppered moth. *The Quarterly review of biology* **51**: 211–44.

Cook L, Saccheri I. **2013**. The peppered moth and industrial melanism: evolution of a natural selection case study. *Heredity* **110**: 207–12.

Costich DE, Meagher TR, Edward J. **1991**. A rapid means of sex identification in *Silene latifolia* by use of flow cytometry. *Plant Molecular Biology Reporter* **9**: 359–370.

Davies MS, Snaydon RW. **1973**. Physiological differences among populations of *Anthoxanthum odoratum* L. collected from the Park Grass Experiment, Rothamsted. II. Response to aluminium. *Journal of Applied Ecology* **10**: 47–55.

Davies MS, Snaydon RW. **1976**. Rapid population differentiation in a mosaic environment III. Measures of selection pressures. *Heredity* **36**: 55–66.

Dolezel J, Greilhuber J, Suda J. **2007**. Estimation of nuclear DNA content in plants using flow cytometry. *Nature protocols* **2**: 2233–44.

Famoso AN, Zhao K, Clark RT, Tung C-WW, Wright MH, Bustamante C, Kochian L V, McCouch SR. **2011**. Genetic architecture of aluminum tolerance in rice (Oryza sativa) determined through genome-wide association analysis and QTL mapping. *PLoS genetics* **7**: e1002221.

Foy CD. **1988**. Plant adaptation to acid, aluminum ‐ toxic soils. *Communications in Soil Science and Plant Analysis* **19**: 959–987.

Freeland JR, Biss P, Silvertown J. **2012**. Contrasting patterns of pollen and seed flow influence the spatial genetic structure of sweet vernal grass (Anthoxanthum odoratum) populations. *The Journal of heredity* **103**: 28–35.

Fujii M, Yokosho K, Yamaji N, Saisho D, Yamane M, Takahashi H, Sato K, Nakazono M, Ma JF. **2012**. Acquisition of aluminium tolerance by modification of a single gene in barley. *Nature communications* **3**: 713.

Fujisawa H, Eguchi S, Ushijima M, Miyata S, Miki Y, Muto T, Matsuura M. **2004**. Genotyping of single nucleotide polymorphism using model-based clustering. *Bioinformatics* **20**: 718–26.

Grime JP, Mowforth MA. **1982**. Variation in genome size - an ecological interpretation. *Nature* **299**: 151.

Huang C-F, Yamaji N, Chen Z, Ma JF. **2011**. A tonoplast-localized half-size ABC transporter is required for internal detoxification of aluminum in rice. *The Plant Journal* **69**: 857-867.

Huang CF, Yamaji N, Mitani N, Yano M, Nagamura Y, Ma JF. **2009**. A bacterial-type ABC transporter is involved in aluminum tolerance in rice. *The Plant Cell* **21**: 655–667.

Kochian L V., Piñeros M a., Hoekenga O a. 2005. The Physiology, Genetics and Molecular Biology of Plant Aluminum Resistance and Toxicity. *Plant and Soil* **274**: 175–195.

Krill AM, Kirst M, Kochian L V, Buckler ES, Hoekenga OA. **2010**. Association and linkage analysis of aluminum tolerance genes in maize. *PLoS ONE* **5**: e9958.

Ma JF. **2007**. Syndrome of aluminum toxicity and diversity of aluminum resistance in higher plants. *International review of cytology* **264**: 225–52.

Magalhaes J V, Garvin DF, Wang Y, Sorrells ME, Klein PE, Schaffert RE, Li L, Kochian L V. **2004**. Comparative mapping of a major aluminum tolerance gene in sorghum and other species in the Poaceae. *Genetics* **167**: 1905–1914.

Magalhaes VJ, Liu J, Guimaraes CT, Lana UGP, Alves VMC, Wang Y-H, Schaffert RE, Hoekenga OA, Pineros MA, Klein PE, *et al.* 2007. A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. *Nature Genetics* **39**: 1156.

McNeilly T, Antonovics J. **1968**. Evolution in closely adjacent plant populations IV. Barriers to gene flow. *Heredity* **23**: 205–218.

Pereira MP, Perez GE, Balbuena ES. **2007**. European Sweet Vernal Grasses (*Anthoxanthum*: Poaceae, Pooideae, Aveneae): A morphometric taxonomical approach. *Systematic Botany* **32**: 43–59.

Pimentel M, Sahuquillo E, Catalán P. **2007**. Genetic diversity and spatial correlation patterns unravel the biogeographical history of the European sweet vernal grasses (*Anthoxanthum* L., Poaceae). *Molecular phylogenetics and evolution* **44**: 667–84.

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. **2000**. Association mapping in structured populations. *The American Journal of Human Genetics* **67**: 170–181.

Raman H, Zhang K, Cakir M, Appels R, Garvin DF, Maron LG, Kochian L V, Moroni JS, Raman R, Imtiaz M, *et al.* 2005. Molecular characterization and mapping of ALMT1, the aluminium-tolerance gene of bread wheat (Triticum aestivum L.). *Genome* **48**: 781–791.

Ryan PR, Delhaize E. **2001**. Aluminium tolerance in plants and the complexing role of organic acids. *TRENDS in Plant Science* **6**.

Sasaki T, Matsumoto H, Yamamoto Y, Ezaki B, Katsuhara M, Ahn SJ, Ryan PR, Delhaize E. **2004**. A wheat gene encoding an aluminum-activated malate transporter. *The Plant Journal* **37**: 645–653.

Silvertown J, Poulton P, Johnston E, Edwards G, Heard M, Biss PM. **2006**. The Park Grass Experiment: its contribution to ecology. *Journal of Ecology* **94**: 801–814.

Silvertown J, Servaes C, Biss P, Macleod D. **2005**. Reinforcement of reproductive isolation between adjacent populations in the Park Grass Experiment. *Heredity* **95**: 198–205.

Smith TB, Bernatchez L. **2008**. Evolutionary change in human-altered environments. *Molecular ecology* **17**: 1–8.

Snaydon RW, Davies MS. **1976**. Rapid population differentiation in a mosaic environment IV. Populations of Anthoxanthum odoratum at sharp boundaries. *Heredity* **37**: 9–25.

Snaydon RW, Davies TM. **1982**. Rapid Divergence of plant populations in response to recent changes in soil conditions. *Evolution* **36**: 289–297.

Thorvaldsdóttir H, Robinson JT, Mesirov JP. **2012**. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics*. online access: April, 2012.

Thrall PH, Young A. **2000**. AUTOTET: a program for analysis of autotetraploid genotypic data. *Journal of Heredity.* report: 348-9.

van't Hof A, Edmonds N, Dalíková M, Marec F, Saccheri J. **2011**. Industrial melanism in British peppered moths has a singular and recent mutational origin. *Science.* **332**: 958–60.

Van Bergen PF, Nott CJ, Bull ID, Poulton PR, Evershed RP. **1998**. Organic geochemical studies of soils from the Rothamsted Classical Experiments--IV. Preliminary results from a study of the effect of soil pH on organic matter decay. *Organic Geochemistry* **29**: 1779–1795.

Wheeler DM. **1995**. Relative aluminium tolerance of ten species of Graminae. *Journal of Plant Nutrition* **18**: 2305–23